Finsler Geometry, Spacetime & Gravity

From Metrizability of Berwald Spaces to Exact Vacuum Solutions in Finsler Gravity

Sjors Heefer

Finsler Geometry, Spacetime & Gravity

Copyright © 2024 by Sjors Heefer

All rights reserved. This dissertation, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the author.

The research documented in this dissertation has been carried out at the Centre for Analysis, Scientific Computing and Applications (CASA) within the Department of Mathematics and Computer Science, Eindhoven University of Technology.

A catalogue record is available from the Eindhoven University of Technology library

ISBN: 978-90-386-6026-4

Printed by ADC Nederland

Finsler Geometry, Spacetime & Gravity

From Metrizability of Berwald Spaces to Exact Vacuum Solutions in Finsler Gravity

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr. S.K. Lenaerts, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op vrijdag 24 mei 2024 om 13:30 uur

door

Sjors Jacco Heefer

geboren te 's-Hertogenbosch

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr. E.R. van den Heuvel
promotor:	prof.dr. L.M.J. Florack
co-promotor:	dr. A. Fuster
leden:	prof.dr. F. Toschi
	prof.dr. G.R. Cavalcanti (Universiteit Utrecht)
	prof.dr. M. Sánchez Caja (Universidad de Granada)
	dr. N. Voicu (Transilvania University of Brasov)
adviseur(s):	dr. C. Pfeifer (University of Bremen)

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

The most beautiful thing we can experience is the mysterious. It is the source of all true art and science.

- A. EINSTEIN

Abstract

This dissertation has been organized into three parts. Part I introduces the mathematical preliminaries. Part II focuses on Berwald spaces, their characterization, and (pseudo-)Riemann metrizability. Finally, Part III deals with exact vacuum solutions to Finsler gravity.

In Part II we first develop a new characterization of Berwald spaces and discuss several applications. One of these is the Berwald condition for (α, β)-metrics, which we apply to several specific cases of interest: Randers metrics, (generalized) m-Kropina metrics, and exponential metrics.

Next, we show, by means of a general argument as well as a simple counterexample, that Szabó's metrization theorem does not generalize to Finsler spaces of indefinite signature. This important theorem states that the affine connection on a positive definite Berwald space (smooth on the slit tangent bundle) can always be understood as the Levi-Civita connection of some Riemannian metric. In short: every such positive definite Berwald space is Riemann metrizable. In other signatures, however, or more generally in situations with less smoothness, we show that the situation is much more complex.

We investigate the class of m-Kropina metrics in detail and obtain several necessary and sufficient conditions for (local) metrizability. Interestingly, local metrizability turns out to be equivalent to the property of having a symmetric (affine) Ricci tensor. The same equivalence holds, trivially, for several other types of metrics, such as Randers metrics. This observation naturally leads us to hypothesize that such an equivalence may hold more generally.

Then, using our characterization of local metrizability, we classify all four-

dimensional Ricci-flat, locally metrizable m-Kropina metrics. The latter class of metrics is of interest in physics since any such metric yields an exact vacuum solution to Pfeifer and Wohlfarth's field equation in Finsler gravity. Some of our results also extend to generalized m-Kropina metrics.

In Part III, we turn our focus completely to exact solutions of the aforementioned field equations in Finsler gravity. We prove that any Finsler metric constructed solely (but arbitrarily) from a vacuum solution in general relativity, α, and a covariantly constant 1-form β, is a vacuum solution in Finsler gravity. This leads to a large class of exact solutions of (α, β)-type. For specific types of (α, β)-metrics, we have stronger results. For Randers metrics, for instance, we classify all solutions of Berwald type and show that they can all be understood as Finslerian pp-wave spacetimes.

We then discuss the physical interpretation and the observational signature of such Finslerian gravitational waves by investigating their effect on interferometric gravitational wave detectors. Remarkably, we come to the conclusion that this effect is indistinguishable from that of a standard gravitational wave in general relativity with the same waveform. Indeed, we compute the expression for the radar distance—the main observable in interferometer experiments—and find that it is identical to its general relativistic counterpart.

Finally, we obtain a solution of unicorn (i.e. Landsberg but not Berwald) type, which is of interest in cosmology. This solution has cosmological symmetry, i.e. is spatially homogeneous and isotropic, and it is additionally conformally flat, with conformal factor depending only on the timelike coordinate. It turns out that, just as in classical Friedmann-Lemaître-Robertson-Walker (FLRW) cosmology, this conformal factor can be interpreted as the scale factor of the universe. Our solution describes a linearly expanding (or contracting) Finslerian universe.

Contents

Abstract **vii**

List of Symbols and Abbreviations **xv**

Introduction **1**

I Preliminaries **9**

1 Fundamentals of Finsler Geometry **13**
1.1 Notation and conventions 13
1.2 Finsler spaces 14
1.3 Geodesics 17
1.4 Finsler tensor fields 18

2 Nonlinear Connections on $\mathcal{A} \subset TM$ **21**
2.1 Horizontal distributions and parallel transport 21
2.2 Curvature, torsion and metric compatibility 24
2.3 The canonical nonlinear connection 26
2.4 The dynamical covariant derivative 29

3 Linear Connections on the Pullback Bundle π^*TM **31**
3.1 Notation and conventions 31
3.2 Torsion and metric-compatibility 32
3.3 The Berwald and Chern-Rund connections 34
3.4 Geodesic deviation 39

4 Berwald, Landsberg and Unicorn Metrics **41**
4.1 Berwald metrics 41
4.2 Landsberg metrics 43
4.3 The unicorn problem 45

II Berwald Spaces and Metrizability **47**

5 Characterization of Berwald Spaces **51**
5.1 The Berwald condition 51
5.2 (α, β)-metrics of Berwald type 54
5.3 Important classes of (α, β)-metrics 58

6 Pseudo-Riemann Metrizability **77**
6.1 Nonmetrizability of Berwald spaces 78
6.2 Generalized m-Kropina spaces with closed null 1-form 80
6.3 Metrizability of m-Kropina spaces 84
6.4 Ricci-flatness of m-Kropina spaces 88

III Vacuum Solutions in Finsler Gravity **95**

7 Finsler Gravity **99**
7.1 Finsler spacetimes 100
7.2 The field equations in vacuum 102

8 Exact Vacuum Solutions of Berwald Type **109**
8.1 (α, β)-metrics 109

8.2 Randers metrics 115
8.3 Modified Randers metrics 121

9 Finsler Gravitational Waves **137**
9.1 Linearized pp-wave solutions 138
9.2 Observational signature 141
9.3 Concluding remarks 150

10 An Exact Cosmological Unicorn Solution **153**
10.1 A modification of Elgendi's unicorn metrics 154
10.2 Exact solutions 162
10.3 A linearly expanding universe 166
10.4 Concluding remarks 168

Discussion and Outlook **171**

Appendices **179**

A Special Pseudo-Riemannian Metrics **179**
A.1 Metrics admitting a closed null 1-form 179
A.2 Metrics admitting a parallel vector field 181
A.3 Some properties of the pp-wave metric 185

B A Basic Linearity Result **187**

C Determinant of an Indefinite (α, β)-Metric **189**

D Additional Proofs **193**
D.1 Metric-compatibility of a nonlinear connection 193
D.2 Generality of generalized m-Kropina metrics 195

List of Publications **203**

Bibliography **205**

Acknowledgments **215**

About the Author **217**

List of Symbols and Abbreviations

M	Smooth manifold.
x^i	Coordinates on the manifold M.
n	Dimension of the manifold M.
TM	Tangent bundle of M.
TM_0	Slit tangent bundle $TM_0 = TM \setminus \{0\}$.
$\mathcal{A}$	Conic subbundle of TM.
$\tilde{\mathcal{A}}$	$= \mathcal{A} \setminus \{\phi' = 0\}$ if $F = \alpha\phi(\beta/\alpha)$ is an (α, β)-metric.
π^*TM	Pull-back bundle of the tangent bundle TM along the canonical projection $\pi : TM \to M$ or its restriction $\pi : \mathcal{A} \to M$ if $\mathcal{A} \subset TM$ is a conic subbundle.
$\mathfrak{X}(M)$	Set of (smooth) vector fields on M.
$\Omega^1(M)$	Set of 1-forms on M.
E	(Total space of a) vector bundle.
$\Omega^1(M, E)$	Set of 1-forms on M with values in a vector bundle E.
$\Gamma(E)$	Sections of a vector bundle E.
$\Gamma(\pi^*TM)$	Set of all Finsler vector fields on M.
$\mathcal{T}^r_s(\pi^*TM)$	Set of all Finsler (k, l)-tensor fields on M.
F	Finsler metric.
L	Finsler Lagrangian.

g_{ij}	Fundamental tensor of a Finsler Lagrangian.
C_{ijk}	Cartan tensor.
N_i^k	Nonlinear connection coefficients.
∂_i	$= \partial/\partial x^i$, partial derivative.
$\bar{\partial}_i$	$= \partial/\partial y^i$, vertical derivative.
δ_i	$= \partial_i - N_i^k \bar{\partial}_k$, horizontal derivative.
$R^k{}_{ij}$	Nonlinear curvature tensor.
$\bar{R}^k{}_{lij}$	Curvature tensor of the canonical linear connection on TM (for Berwald spaces).
Ric	Finsler-Ricci scalar Ric$= R^i_{ij}y^j$.
R_{ij}	Finsler-Ricci tensor $R_{ij} = \frac{1}{2}\bar{\partial}_i\bar{\partial}_j$Ric.
$\bar{R}_{ij}$	Affine Ricci tensor $\bar{R}_{ij} = \bar{R}^k{}_{ikj}$ (for Berwald spaces).
∇	Linear connection.
$\boldsymbol{\nabla}_D$	Dynamical covariant derivative.
Γ^k_{ij}	Christoffel symbols if $E = TM$, or *Horizontal* Christoffel symbols if $E = \pi^*TM$.
$\bar{\Gamma}^k_{ij}$	*Vertical* Christoffel symbols of a linear connection on $E = \pi^*TM$.
$\mathring{\Gamma}^k_{ij}$	Christoffel symbols of the Levi-Civita connection.
$\mathrm{d}x^i\mathrm{d}x^j$	$= \frac{1}{2}(\mathrm{d}x^i \otimes \mathrm{d}x^j + \mathrm{d}x^j \otimes \mathrm{d}x^i)$, symmetrized tensor product.
$T_{(ij)}$	$= \frac{1}{2}(T_{ij} + T_{ij})$, symmetrized tensor components.
$T_{[ij]}$	$= \frac{1}{2}(T_{ij} - T_{ij})$, antisymmetrized tensor components.
a_{ij}	Components of a pseudo-Riemannian metric.
$\mathring{\nabla}$	Levi-Civita connection of a_{ij}.
y^i	Induced coordinates on the fibers of TM.
A	$= a_{ij}y^iy^j$.
ϵ	$\epsilon = \mathrm{sgn}(A)$.
α	$= \sqrt{\lvert A\rvert} = \sqrt{\lvert a_{ij}y^iy^j\rvert}$.
b_i	Components of a 1-form.
β	$= b_iy^i$.
$\lvert b\rvert^2$	$= a_{ij}b^ib^j$.
s	$= \beta/\alpha$.

A^0	Topological interior of the set A.
$\overline{A}$	Topological closure of the set A.
GR	*General Relativity*.
FLRW	*Friedmann-Lemaître-Robertson-Walker*.
iff	*If and only if*.
LHS	*Left-hand side*.
RHS	*Right-hand side*.
TFAE	*The following are equivalent*.
WLOG	*Without loss of generality*.

Introduction

Since its finalization in 1915, Einstein's general theory of relativity (GR) has proven itself as our most successful theory of gravity yet and today it is a cornerstone of modern physics. It has also been clear for a long time, however, that there is no hope of it being the ultimate answer. Indeed, as illustrated by the famous Penrose-Hawking singularity theorems, the theory even predicts its own demise. Besides, while GR is a *classical* theory, the other forces of nature and the known types of matter in the universe all behave according to the laws of quantum mechanics as described in the standard model of particle physics. In order to have a consistent picture of nature it would thus seem that gravity ought to be described quantum mechanically as well.

The 'quantization' of gravity, often dubbed the holy grail of fundamental physics, has proven to be an extraordinary challenge. Even though approaches such as string theory, loop quantum gravity and others have made considerable progress in this direction, it is fair to say that we still do not have a clue which of these approaches, if any, is on the right track.

Nevertheless, many of the fundamental approaches to quantum gravity seem to converge on the idea that local Lorentz invariance may not be fundamental but rather a low-energy approximation (see [1] for a recent review and references therein). Experimental searches for Lorentz invariance violation are actively ongoing as well [1, 2, 3]. Departure from Lorentz invariance has as a direct consequence that Lorentzian geometry—the geometric foundation of GR—is no longer adequate as a description of spacetime in certain regimes. This is where Finsler geometry comes in.

As concisely put by Chern [4], Finsler geometry is just pseudo-Riemannian

geometry without the quadratic restriction. In other words, it is the natural extension of pseudo-Riemannian geometry in which the squared line element $\mathrm{d}s^2$ is not restricted to be quadratic in the coordinate 1-forms $\mathrm{d}x^i$. Historically, the possibility of considering this type of geometry was already discussed by Riemann in his famous habilitation lecture in 1854 [5, 6]. But the first systematic study of such spaces appeared only in Finsler's 1918 PhD thesis [7]. After this, the theory was developed further by mathematicians such as Cartan, Chern, Rund, Matsumoto, and many others.

As the initial development of GR took place between 1905-1915, it is only natural that the theory was formulated on the basis of pseudo-Riemannian geometry, which was already well developed at that time. And when it was found in the decades that followed that the theory matched experiment to extraordinary precision, the obvious conclusion was that the theory and in particular its geometric foundation were simply correct. With the knowledge that we have today, however, and specifically in light of what has been said above, it is natural to consider the possibility that these foundations may need to be revisited. There are compelling reasons to consider Finsler geometry rather than the more restrictive Lorentzian geometry as the fundamental geometric framework underlying the structure of spacetime and gravity.

First of all, while Lorentzian geometry is obviously inadequate in scenarios with a departure from local Lorentz invariance, it turns out that at the classical level, virtually any such scenario can be described in terms of a Finsler geometry on the spacetime manifold [8, 9, 10]; see e.g. [11, 12, 13] for applications to specific phenomenological quantum gravity models.

Second, as suggested already in 1985 by Tavakol and Van den Bergh [14, 15, 16], the axiomatic approach to GR by Ehlers, Pirani, and Schild ('EPS axiomatics') [17] does not single out Lorentzian geometry—as was widely believed for a long time—but is compatible with Finsler geometry as well. This was originally overlooked due to artificially restrictive differentiability assumptions, as recently (2018) pointed out in [18] and then worked out in detail in [19]. Other axiomatic approaches also allow for more general types of geometry, see e.g. [20].

And finally, Finsler geometry provides the most general geometric framework that is compatible with the standard formulation of the clock postulate[1], which states that the time measured by a clock between two events is given by the length of the clock's spacetime trajectory connecting these events; in this case the Finslerian length rather than the pseudo-Riemannian length.

[1]We remark that Weyl geometry, another generalization of Lorentzian geometry, is also compatible with the clock postulate, but in that case the definition of proper time has to be revised [21].

The idea that Finsler geometry might play a role in fundamental physics has been around at least since 1926 when first attempts were made to unify gravity with electromagnetism [22, 23, 24]. Several decades later the possibility of extending GR into the realm of Finsler geometry started being considered [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. But it was only in 2012 that the action-based approach to Finsler gravity was developed by Pfeifer and Wohlfarth [28, 37]. Structurally, this theory is completely analogous to GR, with the only essential difference being that the space of allowed metrics is enlarged from Lorentzian metrics to Finsler metrics. Einstein's field equation is extended into the realm of Finsler geometry by means of Pfeifer and Wohlfarth's field equation, which is derived from the natural Finsler generalization of the Einstein-Hilbert action and reduces to Einstein's field equation in the Lorentzian limit. In the spirit of Chern, we could say that *Finsler gravity is just general relativity without the quadratic restriction*.

Recently (2019), it was discovered [38] that Pfeifer and Wohlfarth's field equation is in fact the variational completion [39] of an earlier proposal for a field equation by Rutz [27], the latter being given simply by the vanishing of the Finsler-Ricci tensor and derived from—or rather motivated by—the geodesic deviation equation. That is to say that, while $R_{\mu\nu} = 0$ cannot be obtained as the Euler-Lagrange expression of an action functional, Pfeifer and Wohlfarth's equation is the unique Euler-Lagrange equation that is *as close to it as possible*, in a well-defined sense [39]. For reference, in the pseudo-Riemannian setting, the variational completion algorithm transforms Einstein's early proposal for the left-hand side of the field equation—the Ricci tensor $R_{\mu\nu}$—into his final and correct expression—the Einstein tensor $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$—that yields not only the correct vacuum equation (which is equivalent to $R_{\mu\nu} = 0$) but also a consistent matter coupling [39]. In 2020 the field equations corresponding to the natural Finsler generalization of the Einstein-Hilbert-Palatini action were investigated as well [40], and for large and important classes of Finsler spaces (Berwald, Landsberg, and weakly Landsberg spaces) these were found to be equivalent to Pfeifer and Wohlfarth's equation. All this puts the latter on firm ground.

Roughly half of this dissertation, Part III, is devoted to the study of exact solutions to Pfeifer and Wohlfarth's field equation for Finsler gravity, to which we will come back momentarily. The other half, Part II, is more fundamental in nature and concerns the characterization and properties of an important class of Finsler spaces, namely those of Berwald type. Apart from their intrinsic mathematical interest, such spaces are very relevant also in the context of Finsler gravity. Indeed, almost all of the solutions to the Finsler gravity field equations discussed in Part III are of this type.

A Berwald space is a Finsler space that admits a (necessarily unique) torsion-free, metric-compatible linear connection on the tangent bundle, akin to the Levi-Civita connection of a pseudo-Riemannian metric. Such spaces may be thought of as being just slightly more general than pseudo-Riemannian manifolds. While various characterizations of Berwald spaces are known [41], we present here a novel characterization in terms of an arbitrary auxiliary pseudo-Riemannian metric and use it to obtain, in particular, a very useful necessary and sufficient Berwald condition for (α, β)-metrics—Finsler metrics constructed from a pseudo-Riemannian metric α and a 1-form β (Chapter 5). We then apply this result to several specific (α, β)-metrics of interest such as Randers metrics, exponential metrics, and (generalized) m-Kropina metrics, obtaining novel results as well as neatly reproducing some well-known ones.

For any Berwald space, one can ask the natural question of whether its canonical torsion-free metric-compatible connection can be understood as the Levi-Civita connection of some auxiliary pseudo-Riemannian metric. This will be referred to as the question of (pseudo-)Riemann metrizability or just metrizability. Simply put: is any Berwald space metrizable? For positive definite Finsler spaces that are smooth on the entire slit tangent bundle the answer, due to Szabo [42], is well-known and affirmative: the affine connection on any such Berwald space is the Levi-Civita connection of some Riemannian metric. However, in alternative signatures like Lorentzian signature, this question remains almost completely unexplored. We demonstrate here that Szabo's metrization theorem does not extend to the general setting (Chapter 6). In particular, it does not hold in Lorentzian signature. For the class of m-Kropina metrics, we analyze the question of metrizability in detail and we obtain precise necessary and sufficient conditions for local metrizability. We also classify all Ricci-flat, locally metrizable m-Kropina spaces. Since any Ricci-flat Berwald space is automatically an exact solution to Pfeifer and Wohlfarth's field equation, this classification has important implications for Finsler gravity.

And that leads us naturally into Part III, which is devoted to the study of exact solutions to Pfeifer and Wohlfarth's field equation for Finsler gravity. Any solution to Einstein's field equation is trivially also a solution in Finsler gravity, but not many exact, properly Finslerian solutions are known as of yet. Prior to the contributions on which this dissertation is based, the only ones known in the literature were the (m-Kropina type) Finsler pp-waves [43] and their generalization as *very general relativity* (VGR) spacetimes [44]. Here we will extend the list considerably.

Most of the solutions that we present and investigate are of Berwald type (Chapter 8). First, we introduce a large class of general Berwald (α, β)-type solu-

tions; we prove that if the pseudo-Riemannian metric α is chosen to be a general relativistic pp-wave and β the corresponding canonical covariantly constant null 1-form, then *any* Finsler metric constructed from these two building blocks will be an exact vacuum solution. This yields a wide range of Finsler metrics that generalize the well-known pp-waves from GR as well as the (m-Kropina type) Finsler pp-waves obtained in [43].

For Randers metrics—the most abundant type of (α, β)-metric—we prove that Pfeifer and Wohlfarth's field equation is in fact equivalent to Rutz's equation, i.e. the vanishing of the Finsler-Ricci tensor. Using this fact, we (locally) classify all exact vacuum solutions of Berwald-Randers type. Completely analogous results hold for what we call the *modified* Randers metric—a small modification of the standard Randers metric that we introduce because of its satisfactory causal properties.

A natural question that arises is how such spacetimes should be interpreted physically, and in particular, whether and how they can be physically distinguished from their general relativistic counterparts, given by the pseudo-Riemannian metric α. To answer this question, we apply a two-fold linearization scheme to our class of (α, β)-metric solutions. We find that the linearized solutions may be interpreted as Finslerian gravitational waves, and we study their observational signature (Chapter 9). More precisely, we ask the question of what would be observed in a gravitational wave interferometer when such a Finslerian gravitational wave passes the earth, and what would be the difference with a classical general relativistic gravitational wave. To this end, we compute the Finslerian radar distance—the main observable measured by interferometers. Remarkably, when interpreted correctly, the result turns out to be completely equivalent to its GR counterpart [45]. In other words, gravitational wave interferometers simply do not have the ability to distinguish this type of Finsler gravitational waves from the standard ones in GR. We discuss the implications of this result.

Finally, in the last chapter, we present a class of exact solutions of unicorn type (Chapter 10). Above, we already introduced Berwald spaces, roughly speaking, as those Finsler spaces that are only 'slightly' more general than pseudo-Riemannian manifolds. One might say that by going up one level in generality, one arrives at the class of Landsberg spaces. Every Berwald space is also Landsberg, but whether or not the opposite is true has been a long-standing open question in the field, provided one adheres to the most strict definition of a Finsler space. Bao has called these non-Berwaldian Landsberg spaces *'[. . .] unicorns, by analogy with those mythical single-horned horse-like creatures for which no confirmed sighting is available,'* [46] and Matsumoto stated in 2003

that they represent the next frontier of Finsler geometry [46]. In the last two decades, some examples of unicorns have been obtained by Asanov [47], Shen [48], and Elgendi [49] by slightly relaxing the definition of a Finsler space. Here we present a new exact vacuum solution to Pfeifer and Wohlfarth's field equation which is a unicorn that falls into one of the classes introduced by Elgendi. This solution has cosmological symmetry, i.e. is spatially homogeneous and isotropic, and it is additionally conformally flat, with conformal factor depending only on the timelike coordinate. We show that, just as in classical Friedmann-Lemaître-Robertson-Walker (FLRW) cosmology, this conformal factor can be interpreted as the scale factor of the universe; we compute this scale factor as a function of cosmological time, and find that it corresponds to a linearly expanding (or contracting) Finslerian universe.

Organization of this dissertation

This dissertation is organized into three parts. The following is a concise summary of their contents.

- In Part I we introduce the necessary mathematical preliminaries (Chapter 1-4).
- In Part II we present our results pertaining to the characterization of Berwald spaces and their metrizability.
  - Chapter 5 covers the characterization of Berwald spaces.
  - Chapter 6 covers metrizability.
- In Part III we present our results concerning exact solutions to Pfeifer and Wohlfarth's field equation in vacuum.
  - Chapter 7 is an introduction to Finsler gravity and reviews the field equation, some of its properties, and its motivation and derivation.
  - Chapter 8 contains our exact solutions that are of Berwald type.
  - Chapter 9 is concerned with the physical interpretation and observational signature of the solutions obtained in Chapter 8.
  - Chapter 10 exhibits our cosmological unicorn solutions.

Moreover, each part has an accompanying introduction that is somewhat more specific than the general introduction given above.

Part I

Preliminaries

Introduction to Part I: Preliminaries

The purpose of Part I is to develop the mathematical prerequisites that are necessary to properly understand Part II and Part III. We assume familiarity with pseudo-Riemannian geometry (see e.g. [50, 51, 52]) and standard concepts in differential geometry such as connections on vector bundles (see e.g. [53, 54, 55]), but we provide a reasonably self-contained exposition of Finsler geometry, starting with the fundamentals in Chapter 1. Here we introduce the definition of a Finsler space, of geodesics, and we introduce the pullback bundle π^*TM and its associated tensor bundles, leading to the notion of a Finsler tensor field.

We then go on to review the topic of nonlinear (Ehresmann) connections in Chapter 2, we introduce the associated notions of curvature, torsion, and metric-compatibility, and we prove the existence and uniqueness of a torsion-free metric-compatible homogeneous nonlinear connection associated with a Finsler metric, generalizing the Levi-Civita connection. We then introduce the Finsler-Ricci curvature and the dynamical covariant derivative, which both appear in the Finsler gravity field equations.

In Chapter 3 we review the theory of linear connections on the pullback bundle π^*TM and in particular their torsion and metric-compatibility, we introduce the Chern-Rund and Berwald connections associated with a Finsler metric, and we discuss geodesic deviation. The latter will be used as motivation for the field equations.

Finally, Chapter 4 discusses Berwald spaces, Landsberg spaces, and unicorns, i.e. non-Berwaldian Landsberg spaces. These notions are very relevant as Part II is centered around Berwald spaces, and many of the solutions to the Finsler gravity field equations discussed in Part III are of this type. The only exception

is our cosmological unicorn solution which, as the name suggests, is of unicorn type.

Our exposition of the material is based on a wide range of literature. For further reference, we refer to [56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 36, 66, 67, 41].

CHAPTER 1

Fundamentals of Finsler Geometry

After briefly outlining some of our notational conventions, we start by introducing Finsler metrics, Finsler Lagrangians, Finsler spaces, and notions such as the Finslerian length of a curve, the fundamental tensor, the Cartan tensor, and Euler's theorem and some of its immediate consequences. We then introduce geodesics as critical points of some 'energy' functional. Although an alternative equivalent definition can be given in terms of autoparallel ('straight') curves of some connection, which is perhaps more intuitive[1], the one given here can be stated without any reference to a connection and hence it will be our starting point. The equivalence with autoparallel curves will be discussed in detail in the next two chapters. Finally, we introduce the pullback bundle π^*TM, its associated tensor bundles, and the notion of a Finsler tensor field.

1.1 Notation and conventions

Not only in this chapter, but throughout the dissertation we will often work in local coordinates. Given a smooth manifold M of dimension n we will always assume that some chart $\varphi : U \subset M \to \mathbb{R}^n$ is provided, and we will effectively

[1] Mathematically, this is merely a matter of taste. In the context of relativity, on the other hand, the definition in terms of autoparallel curves is the physically relevant one. This is because gravity is understood *not* to be a force and hence the trajectories of freely falling test particles should be the straightest possible lines through the curved spacetime geometry, i.e. autoparallels, i.e. geodesics.

identify any $p \in U$ with its image $x = (x^1, \dots, x^n) \coloneqq \varphi(p) \in \mathbb{R}^n$ under φ. If $p \in U$ then each $X_p \in T_pM$ can be written as $X_p = y^i \partial_i\big|_p =: y^i \partial_i$ (we will often be sloppy and suppress the dependence on the point p), where the tangent vectors $\partial_i \coloneqq \frac{\partial}{\partial x_i}$ make up the coordinate basis of T_pM. This decomposition provides natural local coordinates on the tangent bundle TM via the chart

$$\tilde{\varphi} : TU \to \mathbb{R}^n \times \mathbb{R}^n, \qquad \tilde{\varphi}(p, Y) = (\varphi(p), y^1, \dots, y^n) =: (x, y), \tag{1.1}$$

where TU is given by

$$TU = \bigcup_{p \in U} \{p\} \times T_pM \subset TM. \tag{1.2}$$

These local coordinates on TM in turn provide a natural basis $\{\partial_i, \bar{\partial}_i\}_{i=1,\dots,n}$ of its tangent space $T_{(x,y)}TM$ at (x, y), where we define

$$\partial_i \coloneqq \frac{\partial}{\partial x^i}, \qquad \bar{\partial}_i \coloneqq \frac{\partial}{\partial y^i}. \tag{1.3}$$

The space of smooth functions on M will be denoted by $C^\infty(M)$ and the space of smooth vector fields on M by $\mathfrak{X}(M)$. Given a vector bundle $\pi : E \to M$ over M, we denote its space of smooth sections by $\Gamma(E)$.

Finally, we will use the notation $\mathrm{d}x^i \mathrm{d}x^j$ for the symmetrized tensor product of 1-forms, i.e. $\mathrm{d}x^i \mathrm{d}x^j \equiv \frac{1}{2}(\mathrm{d}x^j \otimes \mathrm{d}x^j + \mathrm{d}x^j \otimes \mathrm{d}x^j)$, and whenever we work in Lorentzian signature we will adhere to the sign convention $(-, +, \dots, +)$ unless otherwise specified.

1.2 Finsler spaces

Let M be a smooth manifold. A conic subbundle of TM is an open subset $\mathcal{A} \subset TM$ which is conic in the sense that $(x, \lambda y) \in \mathcal{A}$ for any $(x, y) \in \mathcal{A}$ and $\lambda > 0$, and which satisfies $\pi(\mathcal{A}) = M$, where $\pi : TM \to M$ is the canonical projection of the tangent bundle. The latter property says that the fibers $\mathcal{A}_x$ of $\mathcal{A}$ are nonempty.

Definition 1.2.1. *A Finsler Lagrangian on a conic subbundle $\mathcal{A}$ is a smooth map $L : \mathcal{A} \to \mathbb{R}$ such that*

- *F is positively homogeneous of degree two:*

$$L(x, \lambda y) = \lambda^2 L(x, y)\,, \quad \forall \lambda > 0\,; \tag{1.4}$$

- *The fundamental tensor, with components* $g_{ij} = \bar{\partial}_i \bar{\partial}_j \left(\frac{1}{2} L\right)$*, is nondegenerate.*

In Section 1.4 we will see that g_{ij} is indeed a (Finsler) tensor field.

Any pseudo-Riemannian metric $a = a_{ij} \mathrm{d}x^i \mathrm{d}x^j$ on M induces a quadratic Finsler Lagrangian $L = a_{ij}(x) y^i y^j$ on $\mathcal{A} = TM$ with fundamental tensor $g_{ij} = a_{ij}$. We say that such a Finsler Lagrangian L is pseudo-Riemannian. In fact, it is easy to see that a Finsler Lagrangian is pseudo-Riemannian if and only if (iff) it is quadratic (in y) iff its fundamental tensor has no directional dependence, $g_{ij} = g_{ij}(x)$. In such a situation the theory reduces to pseudo-Riemannian geometry. By extension, we will also say that L is pseudo-Riemannian if there exists a finite cover $\{\mathcal{A}_i\}_i$ of $\mathcal{A}$ such that L is quadratic on each $\mathcal{A}_i$. For example, a Lagrangian such as $L = |a_{ij} y^i y^j|$ on $\mathcal{A} = \{(x, y) \in TM \,:\, a_{ij}(x) y^i y^j \neq 0\}$ will also be called pseudo-Riemannian. If no such cover exists, then we say that L is *properly Finslerian*. It turns out that any Finsler Lagrangian defined on $\mathcal{A} = TM$ must be pseudo-Riemannian. In the properly Finslerian case, the largest possible domain one can hope for is given by $\mathcal{A} = TM_0$, the slit tangent bundle,

$$TM_0 \coloneqq \{(x, y) \in TM \,:\, y \neq 0\}. \tag{1.5}$$

Even though it is the 2-homogeneous Finsler Lagrangian that enters in most (but not all) formulas in Finsler geometry, it is useful to independently define a similar 1-homogeneous object, the Finsler metric.

Definition 1.2.2. *A Finsler metric on a conic subbundle* $\mathcal{A}$ *is a smooth map* $F : \mathcal{A} \to \mathbb{R}$ *such that*

- F *is positively homogeneous of degree one:*

$$F(x, \lambda y) = \lambda F(x, y)\,, \quad \forall \lambda > 0\,; \tag{1.6}$$

- *The fundamental tensor, with components* $g_{ij} = \bar{\partial}_i \bar{\partial}_j \left(\frac{1}{2} F^2\right)$*, is nondegenerate.*

The length of a curve $\gamma : (a, b) \to M$ is then defined as[2]

$$\ell(\gamma) = \int_a^b |F(\gamma(\lambda), \dot{\gamma}(\lambda))| \, \mathrm{d}\lambda, \qquad \dot{\gamma} = \frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}. \tag{1.7}$$

[2]Strictly speaking the length of γ is defined only when $\dot{\gamma}(\lambda) \in \mathcal{A}$ for all $a < \lambda < b$.

The homogeneity condition (1.6) ensures that this length is invariant under (orientation-preserving) reparameterization.

Note that the two definitions 1.2.1 and 1.2.2 are not equivalent. Any Finsler metric has a canonical Finsler Lagrangian $L = F^2$ associated with it and, conversely, from any Finsler Lagrangian on $\mathcal{A}$ one may define a canonical Finsler metric $F = \sqrt{|L|}$ on $\mathcal{A} \cap \{L \neq 0\}$. However, both of these correspondences $F \leftrightarrow L$ are neither injective nor surjective. That said, it is fair to say that Definition 1.2.1 is the more general definition of the two, as in most applications one only cares about L and not (the sign of) F. If so, one could choose to restrict the analysis to Fisler metrics that satisfy $F \geq 0$, which is typically done. In that case the correspondence $F \mapsto L = F^2$ *is* injective and with this identification, nonnegative Finsler metrics form a subclass of Finsler Lagrangians. Nevertheless, we will see several times throughout this dissertation that it can be useful to allow F to vary in sign, which is why we incorporate Definition 1.2.2 as well.

Definition 1.2.3. *A Finsler space or Finsler manifold is a triple* $(M, \mathcal{A}, L)$ *or* $(M, \mathcal{A}, F)$.

We will sometimes omit the specification of $\mathcal{A}$ and say, loosely speaking, that F is a Finsler metric on M and L a Finsler Lagrangian on M. We say that F or L is positive definite if g_{ij} is positive definite, and more generally that F or L has a certain signature if g_{ij} has that signature. We remark that in the positive definite setting, a Finsler space is most often defined as a triple $(M, \mathcal{A}, F)$, where $\mathcal{A} = TM_0$ and F is a Finsler metric in the sense of Definition 1.2.2 with the additional requirement that $F > 0$ and that g_{ij} be positive definite. In other signatures such as Lorentzian signature, however, many important examples have a domain $\mathcal{A}$ which is strictly smaller and cannot be extended to TM_0.

Next, we recall a classical result, known as Euler's theorem for homogeneous functions (stated in a somewhat atypical way). A smooth function $f : U \to \mathbb{R}$ on an open set $U \subset \mathbb{R}^n$ is said to be (positively) homogeneous of degree k or simply k-homogeneous if $f(\lambda y) = \lambda^k f(y)$ for all $\lambda > 0$. If f is a function on a conic subbundle $\mathcal{A} \subset TM$ then, as a matter of speaking, we will say that f is k-homogeneous if it is k-homogeneous in its y-dependence.

Theorem 1.2.4 (Euler's Theorem [56])**.** *Let $f : \mathcal{A} \to \mathbb{R}$ be a smooth, k-homogeneous function on a conic subbundle $\mathcal{A} \subset TM$. Then*

$$y^i \bar{\partial}_i f(x, y) = k f(x, y). \tag{1.8}$$

This result has a central importance in Finsler geometry. For example, since the Finsler Lagrangian L is 2-homogeneous, Euler's theorem implies that it can be

written as

$$L = g_{ij} y^i y^j. \tag{1.9}$$

The Cartan tensor (or Cartan torsion) is defined by its components

$$C_{ijk} = \tfrac{1}{2}\bar{\partial}_i g_{jk} = \tfrac{1}{4}\bar{\partial}_i\bar{\partial}_j\bar{\partial}_k L. \tag{1.10}$$

Note that C_{ijk} is completely symmetric. As a second application of Euler's theorem, the fact that g_{ij} is 0-homogeneous implies that

$$y^i C_{ijk} = y^j C_{ijk} = y^k C_{ijk} = 0. \tag{1.11}$$

The mean Cartan tensor (or torsion) is the trace of the Cartan tensor,

$$C_k = g^{ij} C_{ijk}. \tag{1.12}$$

Rather than taking the Finsler metric F as the fundamental object, one could equivalently start from the fundamental tensor, but then one needs an additional axiom to guarantee that g_{ij} is the Hessian of some scalar function, namely the constraint that the Cartan tensor associated with g_{ij} be fully symmetric. More precisely (see e.g. [65]) we have the following.

A symmetric and nondegenerate collection of smooth, 0*-homogeneous, functions* $g_{ij} : \mathcal{A} \to \mathbb{R}$ *on a conic subbundle* $\mathcal{A}$ *makes up the fundamental tensor of a Finsler Lagrangian if and only if the associated Cartan tensor is completely symmetric.*

1.3 Geodesics

The simplest way to define the notion of geodesics directly from the Finsler function or Lagrangian is via a variational approach. Let $\gamma : [a, b] \to M$ be a curve in a Finsler space such that $\dot{\gamma}(\lambda) \in \mathcal{A}$ for all λ. Then we say that γ is a geodesic if it is a critical point of the 'energy' functional

$$E[\gamma] = \int_a^b L(\gamma(\lambda), \dot{\gamma}(\lambda))\, \mathrm{d}\lambda, \tag{1.13}$$

which is the case iff it is a solution to the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^i} = 0, \tag{1.14}$$

evaluated along $(x, y) = (\gamma(t), \dot{\gamma}(t))$. Using (1.11) it is straightforward to show that Equation (1.14) is equivalent to the *geodesic equation*,

$$\ddot{\gamma}^k + G^k(\gamma, \dot{\gamma}) = 0, \tag{1.15}$$

where the *geodesic spray coefficients*

$$G^k(x, y) = \gamma^k_{ij}(x, y) y^i y^j \tag{1.16}$$

are defined in terms of the so-called *formal Christoffel symbols*[3]

$$\gamma^k_{ij} = \tfrac{1}{2} g^{k\ell} \left(\partial_i g_{\ell j} + \partial_j g_{i\ell} - \partial_\ell g_{ij} \right). \tag{1.17}$$

We will also refer to G^k simply as the geodesic spray. Note the resemblance to the Christoffel symbols of the Levi-Civita connection in Riemannian geometry. The term 'formal' refers to the fact that the γ^k_{ij} are not really the Christoffel symbols of a well-defined connection on TM, except when L is pseudo-Riemannian, in which case they are just the Christoffel symbols of the Levi-Civita connection. In Chapter 2 we will see that geodesics can also be characterized as autoparallel curves of a canonical connection that generalizes the Levi-Civita connection. We also note that geodesics are critical points of the length functional (1.7) as well.

1.4 Finsler tensor fields

Apart from tensor fields on M and tensor fields on $\mathcal{A} \subset TM$ in the standard sense, many of the geometric objects of interest in Finsler geometry are tensor fields on M 'with dependence on both position and direction'. Examples we have seen so far are the fundamental tensor g_{ij} and the Cartan tensor C_{ijk}. For a careful treatment of such tensors, it is useful to introduce the so-called pullback bundle π^*TM. A direction-dependent vector field can then be defined as a section of this bundle, and a direction-dependent tensor field as a section of an associated tensor bundle. We will call such objects *Finsler tensor fields*.

[3] It is a matter of coincidence that both the formal Christoffel symbols and the curve under consideration are represented by the symbol γ here. The formal Christoffel symbols are defined independently of any curve.

This formulation in terms of the pullback bundle will also allow us to easily define covariant derivatives of such tensor fields; both linear covariant derivatives (Chapter 3) as well as the so-called dynamical covariant derivative arising from the canonical nonlinear connection on a Finsler space (Chapter 2).

1.4.1 The pullback bundle π^*TM

Given any vector bundle $\pi : E \to M$ over a smooth manifold M and a smooth map $f : N \to M$, where N is another smooth manifold, one can define the pullback bundle f^*E, which is a smooth vector bundle over N. More precisely, the pullback bundle is the vector bundle $\pi' : f^*E \to N$, where the total space is defined as

$$f^*E = \{(n, e) \in N \times E \mid f(n) = \pi(e)\} \subset N \times E, \tag{1.18}$$

and the projection is defined as $\pi'(n, e) = n$.

Consider now the special case where E is the tangent bundle, $\pi : TM \to M$ the canonical projection, and f its restriction $f = \pi|_{\mathcal{A}} : \mathcal{A} \to M$ to some conic subbundle $\mathcal{A} \subset TM$. Then the construction yields a vector bundle on $\mathcal{A}$, denoted by π^*TM, and to which we will refer as *the* pullback bundle.

1.4.2 Vector fields and tensor fields

As already alluded to above, we shall adopt the following definition.

Definition 1.4.1. *A Finsler vector field (or anisotropic vector field) is a section $X \in \Gamma(\pi^*TM)$ of the pullback bundle.*

Let us check that this definition agrees with our intuitive notion of a Finsler vector field as a vector field with a dependence on both position and direction. By construction, the fiber $\pi^*TM_{(x,y)}$ at (x, y) is canonically isomorphic to the vector space T_xM, so we may think of π^*TM as the vector bundle constructed by erecting a copy of T_xM, the fiber, at each point $(x, y) \in \mathcal{A}$. Given a vector field X on M, one can define a corresponding Finsler vector field $\hat{X}$ by setting $\hat{X}_{(x,y)} = X_x$, using the identification discussed above. If we take for X the local coordinate basis vectors ∂_i, then the corresponding Finsler vector fields $\hat{\partial}_i$ form a basis for the fibers $\pi^*TM_{(x,y)}$. By some abuse of notation, we will drop the hat and denote these simply by ∂_i as well. Finsler vector fields $X \in \Gamma(\pi^*TM)$ can therefore be represented locally as $X = X^i(x, y)\partial_i$, justifying their definition as vector fields with a dependence on both position and direction. As usual, X is smooth if and only if its component functions $X^i(x, y)$ are smooth.

From here the route to defining general Finsler tensor fields is straightforward. Skipping some of the technical details (which can be found in e.g. [57]), the idea is to define the dual bundle $(\pi^*TM)^*$—which is in fact isomorphic to the pullback of the cotangent bundle T^*M along the same map $f = \pi|_{\mathcal{A}} : \mathcal{A} \to M$ that we used earlier—and consider tensor products of π^*TM and $(\pi^*TM)^*$.

Definition 1.4.2. *A Finsler tensor field (or anisotropic tensor field) of type (k, l) is a section of the tensor bundle*

$$\underbrace{\pi^*TM \otimes \cdots \otimes \pi^*TM}_{k \text{ times}} \otimes \underbrace{(\pi^*TM)^* \otimes \cdots \otimes (\pi^*TM)^*}_{l \text{ times}}. \tag{1.19}$$

The space of smooth Finsler (k, l)-tensor fields will be denoted by $\mathcal{T}^k_l(\pi^*TM)$.

Equivalently, a Finsler (k, l)-tensor field T may be regarded as a $C^\infty(\mathcal{A})$-multilinear map

$$T : \underbrace{(\pi^*TM)^* \times \cdots \times (\pi^*TM)^*}_{k \text{ times}} \times \underbrace{\pi^*TM \times \cdots \times \pi^*TM}_{l \text{ times}} \to C^\infty(\mathcal{A}), \tag{1.20}$$

and locally it can be expressed as

$$T = T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l} \partial_{i_1} \otimes \cdots \otimes \partial_{i_k} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_l}, \tag{1.21}$$

where the Finsler 1-forms $\mathrm{d}x^i$ are defined in a way analogous to the Finsler vector fields ∂_i and where the components $T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l}(x, y)$ are functions on $\mathcal{A}$ that behave in the standard way under coordinate transformations, namely

$$\tilde{T}^{i_1 \ldots i_k}{}_{j_1 \ldots j_l} = \frac{\partial \tilde{x}^{i_1}}{\partial x^{m_1}} \cdots \frac{\partial \tilde{x}^{i_k}}{\partial x^{m_k}} \frac{\partial x^{n_1}}{\partial \tilde{x}^{j_1}} \cdots \frac{\partial x^{n_l}}{\partial \tilde{x}^{j_l}} T^{m_1 \ldots m_k}{}_{n_1 \ldots n_l}. \tag{1.22}$$

Finally, by convention, we define a Finsler $(0, 0)$-tensor field to be just a function $f : \mathcal{A} \to \mathbb{R}$. Since we will be dealing with Finsler tensor fields a lot, we will often refer to them simply as tensor fields, and similarly for Finsler vector fields, etc. To distinguish Finsler tensor fields from conventional tensor fields on M, we will sometimes refer to the latter as *classical* tensor fields in case the distinction is important. Under the relevant identifications, a classical tensor field may be regarded as a Finsler tensor field whose components depend only on x and not on y.

CHAPTER 2

Nonlinear Connections on $\mathcal{A} \subset TM$

Nonlinear connections are an essential ingredient in the study of Finsler geometry, as they allow for a generalization of the Levi-Civita connection to Finsler spaces. The name is somewhat misleading, though, as it is perfectly possible for a nonlinear connection to be linear. It would be more accurate to call them *not-necessarily-linear connections*, but we'll stick to the standard terminology. Below we start by briefly recalling the definition of a nonlinear connection in terms of horizontal distributions and the corresponding notion of parallel transport. Then we discuss the notions of curvature, torsion, and compatibility with a Finsler metric, and we prove that, given any Finsler metric there exists a unique homogeneous, torsion-free, and metric-compatible nonlinear connection. We end the chapter by introducing the dynamical covariant derivative on Finsler tensor fields.

2.1 Horizontal distributions and parallel transport

2.1.1 Horizontal distributions

A nonlinear (or Ehresmann) connection on a conic subbundle $\mathcal{A} \subset TM$, later to be identified as the domain of a given Finsler metric, is a smooth decomposition

of the tangent bundle $T\mathcal{A}$ into a horizontal and a vertical subbundle[1],

$$T\mathcal{A} = H\mathcal{A} \oplus V\mathcal{A}, \tag{2.1}$$

where $\oplus$ denotes the Whitney sum of vector bundles. The vertical subbundle $V\mathcal{A}$ is canonically defined on any smooth manifold as

$$V\mathcal{A} = \ker(\mathrm{d}\pi) = \mathrm{span}\left\{\bar{\partial}_i\right\}, \tag{2.2}$$

in terms of the vertical coordinate vector fields (or vertical derivatives) $\bar{\partial}_i$, but in general, there is no preferred choice of the horizontal subbundle. A nonlinear connection is thus simply a specification of this horizontal subbundle $H\mathcal{A}$. In a given chart, the latter can be represented as

$$H\mathcal{A} = \mathrm{span}\left\{\delta_i\right\}, \qquad \delta_i \equiv \partial_i - N^j{}_i \bar{\partial}_j, \tag{2.3}$$

for a collection of smooth functions $N^i_j(x, y)$, $i, j = 1, \ldots, n = \dim M$ on $\mathcal{A}$ (or, strictly speaking, rather on the induced chart on $\mathcal{A}$), known as the *connection coefficients*. The δ_i are called horizontal coordinate vector fields[2] or horizontal derivatives. The connection coefficients have the following transformation behavior under coordinate transformations on M:

$$N^k{}_j = \tilde{N}^l{}_i \frac{\partial x^k}{\partial \tilde{x}^l} \frac{\partial \tilde{x}^i}{\partial x^j} + \frac{\partial x^k}{\partial \tilde{x}^l} \frac{\partial^2 \tilde{x}^l}{\partial x^i \partial x^j} \tilde{y}^i. \tag{2.4}$$

Conversely, any choice of smooth local functions $N^i_j(x, y)$ satisfying (2.4) defines a nonlinear connection.

Once a connection has been specified one can introduce the notions of horizontal and vertical vector fields. These are special types of vector fields on $\mathcal{A}$. We say that a smooth vector field $X \in \mathfrak{X}(\mathcal{A})$ is *horizontal* and write $X \in \mathfrak{X}^h(\mathcal{A})$ if $X_{(x,y)} \in H_{(x,y)}\mathcal{A}$ for all $(x, y) \in \mathcal{A}$. Similarly, we say that a smooth vector field $X \in \mathfrak{X}(\mathcal{A})$ is *vertical* and write $X \in \mathfrak{X}^v(\mathcal{A})$ if $X_{(x,y)} \in V_{(x,y)}\mathcal{A}$ for all $(x, y) \in \mathcal{A}$.

[1] For alternative equivalent definitions, see [40] and references therein.

[2] Note that, strictly speaking, the δ_i are not themselves coordinate vector fields on TM in the usual sense of the word (this is prohibited by the Frobenius theorem unless the curvature of the connection vanishes). The terminology is nevertheless natural here, though, as the δ_i are the horizontal lifts of the coordinate vector fields ∂_i on M.

2.1.2 Parallel transport

A nonlinear connection provides a very general means of defining parallel transport of vectors along curves and it leads in particular to the notion of autoparallel (i.e. 'straight') curves. Parallel transport of a vector field V along a curve γ on M is characterized[3] by the requirement that the rate of change $\dot{V}$ along γ be horizontal, i.e. $\dot{V} \in H\mathcal{A}$. Here we identify V with the curve $t \mapsto (\gamma(t), V(\gamma(t)))$ in $\mathcal{A} \subset TM$ and hence in general we have

$$\dot{V} \equiv \tfrac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t)) = (\dot{\gamma}^1, \dots \dot{\gamma}^n, \dot{V}^1, \dots, \dot{V}^n) = \dot{\gamma}^i \partial_i\big|_V + \dot{V}^i \bar{\partial}_i\big|_V \in TTM. \tag{2.5}$$

We can decompose this as

$$\dot{V}(t) = \dot{\gamma}^i \delta_i\big|_V + \left(N^i_j(\gamma, V)\dot{\gamma}^j + \dot{V}^i\right) \bar{\partial}_i\big|_V, \tag{2.6}$$

in terms of horizontal and vertical coordinate vector fields and hence V is parallel transported along γ iff it satisfies the *parallel transport equation*,

$$\dot{V}^i + N^i_j(\gamma, V)\dot{\gamma}^j = 0\,. \tag{2.7}$$

Notice that parallel transport can be defined only as long as $V(t)$ stays in $\mathcal{A}$. An autoparallel curve, or just autoparallel, is a curve whose tangent vector field is parallel transported along the curve itself. Hence, setting $V = \dot{\gamma}$, a curve γ is an autoparallel if and only if it satisfies the autoparallel equation,

$$\ddot{\gamma}^i + N^i_j(\gamma, \dot{\gamma})\dot{\gamma}^j = 0\,. \tag{2.8}$$

Assuming for the moment that $\mathcal{A} = TM$ it is easy to see from (2.7) that parallel transport is linear in V if and only if N^i_j is linear in V. Therefore we say that a connection is linear if its connection components are linear in the tangent space coordinates, i.e. $N^k_i(x, y) = \Gamma^k_{ij}(x)y^j$. Even if $\mathcal{A}$ is not a linear space, the condition $N^k_i(x, y) = \Gamma^k_{ij}(x)y^j$ for all $(x, y) \in \mathcal{A}$ allows one to canonically extend the nonlinear connection to $\mathcal{A} = TM$, so it still makes sense in this case to say the connection is linear, hence we will do so. If the connection is linear then, as a result of (2.4), the functions Γ^k_{ij} are the Christoffel symbols[4] of a linear

[3] Here we follow e.g. [57]. We note that some authors (see e.g. [58]) choose to define parallel transport in a different way, namely by requiring *a priori* that parallel transport should be linear, which then leads to the alternative parallel transport equation $\dot{V}^i + N^i_j(\gamma, \dot{\gamma})V^j = 0$, where $\dot{\gamma}$ and V are interchanged in the second term with respect to (2.7). In certain specific situations, this description has its advantages, but in general, it seems less natural to us.

[4] Sometimes the term *Christoffel symbols* is reserved for the connection coefficients of the Levi-Civita connection. For us, the term will apply to the connection coefficients of any linear

(Koszul) connection on TM, i.e. an affine connection on M, in the standard sense. Thus, essentially, a linear nonlinear connection (excuse the terminology) is nothing but a linear Koszul connection and we will indeed identify the two whenever no confusion is possible. In such a case, (2.8) reduces to the familiar autoparallel equation of the corresponding Koszul connection,

$$\ddot{\gamma}^i + \Gamma^i_{jk}(\gamma)\dot{\gamma}^j\dot{\gamma}^k = 0\,. \tag{2.9}$$

Similarly, we say that a nonlinear connection is positively homogeneous of degree 1, or simply *homogeneous*, if $N^k_i(x, \lambda y) = \lambda N^k_i(x, y)$ for all $\lambda > 0$. This definition makes sense since $\mathcal{A}$ is conic, by definition.

2.2 Curvature, torsion and metric compatibility

It is the goal of this section to introduce the notions of curvature, torsion, and metric-compatibility for nonlinear connections on $\mathcal{A} \subset TM$. To this end, we first recall the definitions of the horizontal and vertical lift of a vector field. To any (classical) vector field on the base manifold M, we can assign both a horizontal and a vertical (classical) vector field on $\mathcal{A}$. The horizontal lift X^h of a vector field $X \in \mathfrak{X}(M)$ is the horizontal vector field $X^h \in \mathfrak{X}(\mathcal{A})$ defined by linear extension of $\partial_i \mapsto \delta_i$, i.e. if $X = X^i\partial_i$ then

$$X^h = X^i\delta_i, \tag{2.10}$$

where the components X^i are now interpreted as functions of x and y; technically we should write $X^h = (X^i \circ \pi)\delta_i$ but we will usually omit the projection π. Similarly, the vertical lift $X^v \in \mathfrak{X}(\mathcal{A})$ of X is defined by linear extension of $\partial_i \mapsto \bar{\partial}_i$, i.e.

$$X^v = X^i\bar{\partial}_i. \tag{2.11}$$

2.2.1 Curvature and torsion

The horizontal and vertical lifts of vector fields on M provide an intuitive way of understanding the curvature and torsion of a nonlinear connection. The idea is that the curvature measures the failure of two horizontal (coordinate) vector fields to commute, while the torsion measures the failure of a horizontal and a vertical (coordinate) vector field to commute. In other words, we are interested in

connection, to distinguish it from the *connection coefficients* of a nonlinear connection.

$[\delta_i, \delta_j]$ and $[\delta_i, \bar{\partial}_j]$, or more generally, $[X^h, Y^h]$ and $[X^h, Y^v]$ for any $X, Y \in \mathfrak{X}(M)$. The latter two commutators are not tensorial, however. In order to make them tensorial, that is, $C^\infty(M)$-multilinear, we need to modify the expressions in the following way,

$$R(X,Y) = -\left([X^h, Y^h] - [X,Y]^h\right), \tag{2.12}$$

$$T(X,Y) = [X^h, Y^v] - [Y^h, X^v] - [X,Y]^v. \tag{2.13}$$

The tensor fields R and T are called the *curvature* and the *torsion* of the nonlinear connection, respectively[5] (see e.g. [57]). They are smooth $C^\infty(M)$-bilinear maps $\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}^v(\mathcal{A})$. It is easily deduced that their coordinate expressions are given by

$$R(\partial_i, \partial_j) = -[\delta_i, \delta_j] =: R^k_{ij}\bar{\partial}_k, \qquad R^k_{ij} = \delta_i N^k_j - \delta_j N^k_i, \tag{2.14}$$

$$T(\partial_i, \partial_j) = [\delta_i, \bar{\partial}_j] - [\delta_j, \bar{\partial}_k] =: T^k_{ij}\bar{\partial}_k, \qquad T^k_{ij} = \bar{\partial}_j N^k_i - \bar{\partial}_i N^k_j. \tag{2.15}$$

If the connection is linear, $N^k_i = \Gamma^k_{ij}(x)y^j$, then a straightforward calculation reveals that the expressions above reduce to

$$R^k_{ij} = \bar{R}^k{}_{\ell ij}(x)y^\ell, \qquad T^k_{ij} = \Gamma^k_{ij}(x) - \Gamma^k_{ji}(x), \tag{2.16}$$

in terms of the curvature tensor of the associated affine connection and its Ricci tensor,

$$\bar{R}^k{}_{\ell ij} = \partial_i\Gamma^k_{j\ell} - \partial_j\Gamma^k_{i\ell} + \Gamma^k_{im}\Gamma^m_{j\ell} - \Gamma^k_{jm}\Gamma^m_{i\ell}, \qquad \bar{R}_{lk} = \bar{R}^i{}_{lik}, \tag{2.17}$$

respectively. In other words, we recover the standard expressions for the components of the curvature tensor and torsion tensor of a linear connection.

2.2.2 Metric-compatibility

We say that a nonlinear connection is compatible with a Finsler metric, or simply metric-compatible, if the Finsler Lagrangian L is invariant under parallel transport. Or in other words, if the function $t \mapsto L(\gamma(t), v(t))$ is constant whenever v is a vector field that is parallel along a curve $\gamma : t \mapsto \gamma(t)$. Sometimes this

[5] The overall minus sign in the definition of the curvature tensor is chosen so as to guarantee that the curvature tensor reduces to the standard curvature tensor for linear connections in the special case that the connection is linear.

property is also called *weak* metric-compatibility[6].

Proposition 2.2.1. *A nonlinear connection is metric-compatible if and only if the Finsler Lagrangian L is horizontally constant, i.e. $\delta_i L = 0$.*

Proof. For parallel transport of a vector v along a curve $\gamma(t)$ in M we have $\dot{v}^i + N^i_j(\gamma, v)\dot{\gamma}^j = 0$, by (2.7), and hence

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} L(\gamma(t), v(t)) &= \dot{\gamma}^i \partial_i L + \dot{v}^i \bar{\partial}_i L && (2.18)\\ &= \dot{\gamma}^i \partial_i L - N^i_j(\gamma, v)\dot{\gamma}^j \bar{\partial}_i L && (2.19)\\ &= \dot{\gamma}^j \left(\partial_j L - N^i_j(\gamma, v) \bar{\partial}_i L \right) && (2.20)\\ &= \dot{\gamma}^j \delta_j L\big|_{(\gamma(t), v(t))}. && (2.21) \end{aligned}$$

This vanishes for all γ, v if and only if $\delta_j L = 0$. □

In some of the literature on Finsler geometry, an alternative definition of metric-compatibility is used, namely that $\boldsymbol{\nabla}_D g = 0$ in terms of the dynamical covariant derivative $\boldsymbol{\nabla}_D$ to be introduced shortly in Section 2.4. The following result shows that this definition is equivalent to ours, provided that the nonlinear connection is assumed to be torsion-free, which it always is in the context of Finsler geometry (cf. Section 2.3).

Proposition 2.2.2. *Let F be a Finsler metric and N a torsion-free, homogeneous nonlinear connection on M. Then*

$$\delta_i L = 0 \qquad \Longleftrightarrow \qquad \boldsymbol{\nabla}_D g = 0. \tag{2.22}$$

Since the proof is somewhat lengthy and the result will not be used anywhere else in this dissertation, the proof can be found in Appendix D.1.

2.3 The canonical nonlinear connection

Just as one can assign a canonical linear connection to any pseudo-Riemannian manifold, the Levi-Civita connection, one can similarly assign a canonical connection to any Finsler manifold, if one is willing to give up linearity. Geodesics, as we defined them in Section 1.3, can then be understood as autoparallels of this

[6] In contrast to strong-metric compatibility, which is the property that the fundamental tensor is horizontally constant with respect to the Berwald derivative ∇^B induced by the nonlinear connection, see e.g. [66].

canonical connection, and the notions of nonlinear curvature and Ricci curvature of a Finsler manifold can be defined.

2.3.1 The fundamental lemma of Finsler geometry

The nonlinear connection established by the following theorem is called the *canonical nonlinear connection* or the *Cartan nonlinear connection*. We give a simple proof in local coordinates. The first intrinsic, coordinate-free proof was given by Grifone [68], see also [57].

Theorem 2.3.1 (Fundamental lemma of Finsler geometry)**.** *Given a Finsler Lagrangian L defined on $\mathcal{A} \subset TM$, there is a unique homogeneous nonlinear connection on $\mathcal{A}$ that is torsion-free and compatible with L. Its connection coefficients are given by*

$$N_i^j = \tfrac{1}{4}\bar{\partial}_i \left[g^{jk} \left(y^m \partial_m \bar{\partial}_k L - \partial_k L \right) \right]. \tag{2.23}$$

Proof. The three desired properties can be stated as follows:

(i) $\partial_i L - N_i^j \bar{\partial}_j L = 0$ (metric-compatibility);

(ii) $\bar{\partial}_j N_i^k = \bar{\partial}_i N_j^k$ (torsion-freeness);

(iii) $N_i^j(x, \lambda y) = \lambda N_i^j(x, y)$, for all $\lambda > 0$ (homogeneity).

To prove uniqueness, we have to show that (i)-(iii) imply (2.23). We begin by differentiating (i) with respect to y^k and multiplying by y^i and then in the middle term we apply (ii) followed by (iii) in the form of Euler's theorem:

$$\begin{aligned} 0 &= y^i \left(\partial_i \bar{\partial}_k L - \bar{\partial}_k N_i^j \bar{\partial}_j L - N_i^j \bar{\partial}_j \bar{\partial}_k L \right) && (2.24)\\ &= y^i \partial_i \bar{\partial}_k L - y^i \bar{\partial}_i N_k^j \bar{\partial}_j L - y^i N_i^j \bar{\partial}_j \bar{\partial}_k L && (2.25)\\ &= y^i \partial_i \bar{\partial}_k L - N_k^j \bar{\partial}_j L - y^i N_i^j \bar{\partial}_j \bar{\partial}_k L. && (2.26) \end{aligned}$$

Next, we bring the rightmost term to the left-hand side (LHS), which we rewrite in terms of the fundamental tensor, and we apply (i) to the right-hand side (RHS). This yields

$$2 y^i N_i^j g_{jk} = y^i \partial_i \bar{\partial}_k L - \partial_k L \tag{2.27}$$

Because g_{ij} is nondegenerate, this is equivalent to

$$2y^i N_i^j = g^{jk}\left(y^i \partial_i \bar{\partial}_k L - \partial_k L\right). \tag{2.28}$$

The last step is to differentiate both sides with respect to y^l and apply (ii) and Euler's theorem once more. The result, after renaming indices, is (2.23). This proves uniqueness, and also existence in a local chart, since (2.23) satisfies all three properties (i)-(iii). Global existence can be verified by checking that the obtained formula N_i^j has the appropriate transformation behavior, namely

$$\tilde{N}^k{}_j = N^l{}_i \frac{\partial \tilde{x}^k}{\partial x^l}\frac{\partial x^i}{\partial \tilde{x}^j} + \frac{\partial \tilde{x}^k}{\partial x^l}\frac{\partial^2 x^l}{\partial \tilde{x}^i \partial \tilde{x}^j}\,\tilde{y}^i. \tag{2.29}$$

This is a little tedious but in principle straightforward. □

One can easily check that if $L = g_{ij}(x)y^i y^j$ arises from a pseudo-Riemannian metric g_{ij} then the canonical connection reduces to the Levi-Civita connection (having Christoffel symbols Γ^k_{ij}) of g_{ij}, in the sense that

$$N_i^j = \Gamma^j_{im} y^m, \qquad \Gamma^j_{mi} = \tfrac{1}{2} g^{jk}\left(\partial_i g_{km} + \partial_m g_{ki} - \partial_k g_{mi}\right). \tag{2.30}$$

2.3.2 Geodesics and autoparallels

For a general Finsler space, by writing out the term in square brackets in the expression (2.23) for the nonlinear connection coefficients and using that $L = g_{ij}(x,y)y^i y^j$, we find that

$$g^{jk}\left(y^m \partial_m \bar{\partial}_k L - \partial_k L\right) = g^{jk}\left(2\partial_\ell g_{ik} - \partial_k g_{i\ell}\right) y^i y^\ell = 2\gamma^j_{i\ell} y^i y^\ell = 2G^j, \tag{2.31}$$

in terms of the geodesic spray coefficients defined in (1.16). The canonical nonlinear connection and the geodesic spray of a Finsler metric are thus related in the following way:

$$N_i^j = \tfrac{1}{2}\bar{\partial}_i G^j, \qquad G^j = N_i^j y^i, \tag{2.32}$$

where the latter equation follows by multiplying the former by y^i and applying Euler's theorem. Moreover, (2.31) yields an expression for the spray directly in terms of the Finsler Lagrangian and fundamental tensor

$$G^j = \tfrac{1}{2} g^{jk}\left(y^m \partial_m \bar{\partial}_k L - \partial_k L\right). \tag{2.33}$$

Computationally, (2.33) often provides the most efficient way of computing the spray coefficients directly from an explicitly given Finsler metric. Finally, by combining (2.32), (1.15), and (2.8) we obtain the following important result.

Proposition 2.3.2. *The geodesics of a Finsler metric coincide with the autoparallels of its canonical nonlinear connection.*

2.3.3 Curvatures on a Finsler space

Given a Finsler metric or Lagrangian, we define its *nonlinear curvature* as the curvature (2.14) of its canonical nonlinear connection. We further define

$$\mathrm{Ric} = R^i{}_{ij} y^j, \qquad R_{ij} = \tfrac{1}{2}\bar{\partial}_i \bar{\partial}_j \mathrm{Ric}. \tag{2.34}$$

In the literature, R_{ij} is often just referred to as the Ricci tensor, but we will give it the name *Finsler-Ricci tensor* to distinguish it from the *affine Ricci tensor* that will be introduced later for Berwald spaces. Similarly, we will refer to Ric as the *Finsler-Ricci scalar*. From homogeneity and Euler's theorem, it follows that the second relation in (2.34) can be inverted as

$$\mathrm{Ric} = R_{ij} y^i y^j. \tag{2.35}$$

Ric and R_{ij} thus contain the same information, and in particular, we have $\mathrm{Ric} = 0$ if and only if $R_{ij} = 0$.

2.4 The dynamical covariant derivative

A nonlinear connection induces a so-called *dynamical covariant derivative* $\boldsymbol{\nabla}_D$ that acts on functions $f \in C^\infty(\mathcal{A})$ and Finsler vector fields $X \in \Gamma(\pi^* TM)$ as

$$\boldsymbol{\nabla}_D f = y^i \delta_i f, \tag{2.36}$$

$$\boldsymbol{\nabla}_D X = (\boldsymbol{\nabla}_D X)^j \, \partial_i, \qquad (\boldsymbol{\nabla}_D X)^j = y^i \delta_i X^j + N^j_i X^i, \tag{2.37}$$

respectively, where $X = X^i \partial_i$ in local coordinates. With this definition, $\boldsymbol{\nabla}_D$ maps functions to functions and Finsler vector fields to Finsler vector fields. As per standard convention, we will usually omit the brackets in (2.37) and write

$$\boldsymbol{\nabla}_D X^j \equiv (\boldsymbol{\nabla}_D X)^j \,. \tag{2.38}$$

We stress that the LHS of (2.38) is not to be interpreted as the dynamical covariant derivative of the component function X^j, but always as the j^{th} component of the vector field $\boldsymbol{\nabla}_D X$. We will adhere to this convention for all expressions of this type, so that no confusion should be possible as to what is meant.

By requiring certain natural properties such as the Leibniz rule, $\boldsymbol{\nabla}_D$ extends uniquely to a dynamical covariant derivative acting on general Finsler tensor fields in the usual way, e.g.

$$\boldsymbol{\nabla}_D g_{ij} \equiv (\boldsymbol{\nabla}_D g)_{ij} = y^k \delta_k g_{ij} - N_i^k g_{kj} - N_j^k g_{ik}. \tag{2.39}$$

The generalization to higher-order tensor fields should be clear from this example. Given a tensor field $T \in \mathcal{T}_s^r(\pi^* TM)$, it follows from the transformation behavior of the components of $\boldsymbol{\nabla}_D T$ that the latter is again a (k,l)-tensor field. In other words, $\boldsymbol{\nabla}_D$ may be thought of as a map

$$\boldsymbol{\nabla}_D : \mathcal{T}_s^r(\pi^* TM) \to \mathcal{T}_s^r(\pi^* TM). \tag{2.40}$$

Lemma 2.4.1. *For any $f \in C^\infty(TM)$ we have*

$$\boldsymbol{\nabla}_D f = \dot{\gamma}^k \delta_k f = \tfrac{\mathrm{d}}{\mathrm{d}t} f \tag{2.41}$$

along any geodesic $\gamma(t)$.

Proof. Along an arbitrary curve γ we have

$$\left(\tfrac{\mathrm{d}}{\mathrm{d}t} f\right)(\gamma, \dot{\gamma}) = \dot{\gamma}^i \partial_i f + \ddot{\gamma}^i \bar{\partial}_i f \tag{2.42}$$

$$= \dot{\gamma}^i \left(\delta_i f + N_i^k \bar{\partial}_k f\right) + \ddot{\gamma}^i \bar{\partial}_i f \tag{2.43}$$

$$= \dot{\gamma}^i \delta_i f + \left(\ddot{\gamma}^k + \dot{\gamma}^i N_i^k\right) \bar{\partial}_k f \tag{2.44}$$

$$= \boldsymbol{\nabla}_D f + \left(\ddot{\gamma}^k + \dot{\gamma}^i N_i^k\right) \bar{\partial}_k f. \tag{2.45}$$

If γ is a geodesic then the term in parentheses vanishes and the result follows. □

We will see in the Chapter 3 that this result can be extended when N_i^k is the canonical nonlinear connection of a Finsler space, such that even for tensors in general, the dynamical covariant can be interpreted as the linear covariant derivative along geodesics with respect to either of the linear Berwald or Chern-Rund connection.

CHAPTER 3

Linear Connections on the Pullback Bundle π^*TM

The canonical nonlinear connection corresponding to a Finsler metric, discussed in Chapter 2, leads to a characterization of geodesics as autoparallel curves of this connection, and it allows us to differentiate Finsler tensor fields by means of the dynamical covariant derivative. Despite its name, the latter is not a covariant derivative in the usual sense, essentially because it does not allow for the differentiation along different directions. For various purposes, it is useful to also introduce a covariant derivative which *can* be used to do just that; and the way to do this is to specify a linear (Koszul) connection on the pullback bundle π^*TM (cf. Section 1.4.1). In contrast to the situation for nonlinear connections, the choice of such a linear connection—even in the presence of a Finsler metric—is in general not unique[1]. We will focus here mostly on the Berwald and Chern-Rund connections.

3.1 Notation and conventions

Recall that a linear (Koszul) connection on a smooth vector bundle E over M is an $\mathbb{R}$-linear map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E), \qquad (X, s) \mapsto \nabla_X s, \tag{3.1}$$

[1] We remark, however, that in the alternative framework of [69], one can single out a unique so-called (linear) 'anisotropic connection', which serves essentially the same purpose.

such that $\nabla_{fX}s = f\nabla_X s$ and $\nabla_X(fs) = (Xf)s + f\nabla_X s$ for any $f \in C^\infty(M)$. Consider now the case where $E = \pi^*TM$. Since the base manifold is the conic subbundle $\mathcal{A} \subset TM$, a linear connection on π^*TM is a map

$$\nabla : \mathfrak{X}(\mathcal{A}) \times \Gamma(\pi^*TM) \to \Gamma(\pi^*TM), \tag{3.2}$$

so in a local coordinate basis $\partial_i \in \mathfrak{X}(M)$ of TM, with corresponding basis $\partial_i \in \Gamma(\pi^*TM)$ of π^*TM and induced coordinate basis $\partial_i, \bar{\partial}_j \in \mathfrak{X}(\mathcal{A})$ of $T\mathcal{A}$, we have, by linearity,

$$\nabla_{\partial_i}\partial_j = \tilde{\Gamma}^k_{ij}\partial_k, \qquad \nabla_{\bar{\partial}_i}\partial_j = \bar{\Gamma}^k_{ij}\partial_k, \tag{3.3}$$

for local functions $\tilde{\Gamma}^k_{ij}, \bar{\Gamma}^k_{ij} \in C^\infty(\mathcal{A})$, called the Christoffel symbols. We have written a tilde above the first set of Christoffel symbols because these are not the ones that we will use. Indeed, if there is a nonlinear connection defined on $\mathcal{A}$ then we will usually employ the corresponding horizontal and vertical basis vectors $\delta_i, \bar{\partial}_i$ of $T\mathcal{A}$, thus defining the Christoffel symbols

$$\nabla_{\delta_i}\partial_j = \Gamma^k_{ij}\partial_k, \qquad \nabla_{\bar{\partial}_i}\partial_j = \bar{\Gamma}^k_{ij}\partial_k. \tag{3.4}$$

From these definitions, it follows that

$$\nabla_{\delta_i}X^j \equiv (\nabla_{\delta_i}X)^j = \delta_i X^j + \Gamma^j_{ik}X^k \tag{3.5}$$

$$\nabla_{\bar{\partial}_i}X^j \equiv \left(\nabla_{\bar{\partial}_i}X\right)^j = \bar{\partial}_i X^j + \bar{\Gamma}^j_{ik}X^k. \tag{3.6}$$

As usual, a Koszul connection on π^*TM as defined above extends uniquely to a connection on all tensor bundles of π^*TM, and one obtains analogous formulas for the horizontal and vertical covariant derivatives of arbitrary Finsler tensors.

3.2 Torsion and metric-compatibility

3.2.1 Torsion

Roughly following [67] (and see also [36] for a slightly different but equivalent approach), given a morphism $\rho : T\mathcal{A} \to \pi^*TM$ of vector bundles over $\mathcal{A}$, or equivalently a (generalized) soldering form[2] $\omega \in \Omega(\mathcal{A}, \pi^*TM)$, one can define a torsion tensor $T^{(\rho)} : \mathfrak{X}(\mathcal{A}) \times \mathfrak{X}(\mathcal{A}) \to \Gamma(\pi^*TM)$ of a linear connection ∇ on

[2]Sometimes a soldering form on a vector bundle E over M is defined as an E-valued 1-form $\omega \in \Omega(M, E)$ with the additional property that for each x, $\omega_x : T_xM \to E_x$ is a linear isomorphism. For our purposes the latter property is not necessary, hence the word *generalized*.

π^*TM by slightly modifying the formula for the torsion of an affine connection in the following natural way:

$$T^{(\rho)}(\tilde{X},\tilde{Y}) = \nabla_{\tilde{X}}\rho(\tilde{Y}) - \nabla_{\tilde{Y}}\rho(\tilde{X}) - \rho([\tilde{X},\tilde{Y}]), \tag{3.7}$$

where $\rho(\tilde{Y})$ is defined as the Finsler vector field such that $\rho(\tilde{Y})_{(x,y)} = \rho(\tilde{Y}_{(x,y)})$. Note, however, that for different morphisms ρ, this leads to different notions of torsion. In our case two morphisms are natural. First, let ρ be given by linear extension of $\rho(\delta_i) = \partial_i$ and $\rho(\bar{\partial}_i) = 0$. The corresponding torsion will be called the *horizontal torsion* and denoted by T_{hor}. It decomposes as

$$T_{\text{hor}}(\delta_i,\delta_j) = \left(\Gamma^k_{ij} - \Gamma^k_{ji}\right)\partial_k, \tag{3.8}$$

$$T_{\text{hor}}(\delta_i,\bar{\partial}_j) = -\bar{\Gamma}^k_{ji}\partial_k, \tag{3.9}$$

$$T_{\text{hor}}(\bar{\partial}_i,\bar{\partial}_j) = 0. \tag{3.10}$$

Next, let ρ be given by linear extension of $\rho(\delta_i) = 0$ and $\rho(\bar{\partial}_i) = \partial_i$. The corresponding torsion will be called the *vertical torsion* and denoted by T_{ver}. It decomposes as

$$T_{\text{ver}}(\delta_i,\delta_j) = R^k_{ij}\partial_k, \tag{3.11}$$

$$T_{\text{ver}}(\delta_i,\bar{\partial}_j) = \left(\Gamma^k_{ij} - \bar{\partial}_j N^k_i\right)\partial_k, \tag{3.12}$$

$$T_{\text{ver}}(\bar{\partial}_i,\bar{\partial}_j) = \left(\bar{\Gamma}^k_{ij} - \bar{\Gamma}^k_{ji}\right)\partial_k, \tag{3.13}$$

where R^k_{ij} is the curvature of the nonlinear connection. To avoid confusion we remark that some authors consider the horizontal T_{hor} to be *the* torsion of the linear connection and do not mention the vertical torsion T_{ver} at all. Under that terminology the well-known Chern-Rund and Berwald connections in Finsler geometry, to be defined below, would be torsion-free. Their vertical torsions do in general not all vanish, though.

3.2.2 Metric compatibility

A linear connection ∇ on π^*TM is said to be compatible with a Finsler metric F, or just *metric-compatible*, if the fundamental tensor is covariantly constant with respect to ∇, i.e. $\nabla g = 0$, that is $\nabla_{\delta_i} g_{jk} = \nabla_{\bar{\partial}_i} g_{jk} = 0$. One might hope that it were possible to associate a torsion-free (or at least horizontally torsion-free), metric-compatible linear connection on π^*TM with any Finsler metric, in analogy with the Levi-Civita connection. This is not possible, however, unless F

is pseudo-Riemannian. A slightly weaker notion of metric-compatibility, to which we will refer as *horizontal* metric-compatibility, is the property that $\nabla_{\delta_i} g_{jk} = 0$. In this case, the fundamental tensor is covariantly constant in all horizontal directions, but not necessarily in the vertical directions. Similarly, we may define vertical metric compatibility by the property that $\nabla_{\bar{\partial}_i} g_{jk} = 0$.

3.3 Berwald and Chern-Rund connections

We are now ready to introduce two of the much-used linear (Koszul) connections in Finsler geometry, the *Berwald connection* and the *Chern-Rund connection*. While the latter is always defined in the context of a given Finsler metric, the Berwald connection can be defined for any nonlinear connection N_i^k, without even specifying a Finsler metric. But by taking N_i^k to be the canonical torsion-free, metric-compatible, homogeneous nonlinear connection associated with a Finsler metric F, we obtain the Berwald connection associated with F. Other important linear connections in Finsler geometry are the linear Cartan connection and the Hashiguchi connection. We will mention these only briefly below.

3.3.1 The Berwald connection

Any homogeneous nonlinear connection on $\mathcal{A} \subset TM$ induces a canonical linear connection ∇^B on π^*TM, known as the *Berwald connection*, whose Christoffel symbols are given by

$$ {}^B\Gamma_{ij}^k = \bar{\partial}_j N_i^k, \qquad {}^B\bar{\Gamma}_{ij}^k = 0. \tag{3.14} $$

By Euler's theorem, this implies that the nonlinear connection coefficients can be obtained from the Christoffel symbols of the Berwald connection as

$$ N_i^k = {}^B\Gamma_{ij}^k y^j. \tag{3.15} $$

Taking into account (2.32) it follows that the Berwald Christoffel symbols can be expressed as

$$ {}^B\Gamma_{ij}^k = \tfrac{1}{2}\bar{\partial}_i \bar{\partial}_j G^k \tag{3.16} $$

in terms of the geodesic spray, and conversely, that

$$ G^k = {}^B\Gamma_{ij}^k y^i y^j. \tag{3.17} $$

From the expressions (3.8)–(3.13) for the torsions and the definition of the Berwald connection (3.14) it is clear that one can characterize the Berwald connection as the unique linear connection for which both mixed torsions vanish, i.e. $T_{\mathrm{hor}}(\delta_i, \bar{\partial}_j) = T_{\mathrm{ver}}(\delta_i, \bar{\partial}_j) = 0$. Note that these requirements imply that the remaining torsions of the Berwald connection vanish identically as well, except for the parts $T_{\mathrm{hor}}(\delta_i, \delta_j) = T^k_{ij}\partial_k$ and $T_{\mathrm{ver}}(\delta_i, \delta_j) = R^k_{ij}\partial_k$, which are given by the torsion and the curvature of the nonlinear connection, respectively. It is therefore fair to say that the Berwald connection is the closest to a completely torsion-free linear connection on π^*TM as one can find in the presence of a given nonlinear connection.

If N^k_i is the canonical nonlinear connection on a Finsler space then the corresponding Berwald connection is called the Berwald connection associated with the Finsler space, or simply *the* Berwald connection. In this case (and more generally in any situation where the nonlinear connection is torsion-free) the Berwald covariant derivative is related to the dynamical covariant derivative $\boldsymbol{\nabla}_D$ in the following way.

Proposition 3.3.1. *The dynamical covariant derivative $\boldsymbol{\nabla}_D$ and the Berwald connection ∇^B, both corresponding to the same homogeneous, torsion-free nonlinear connection, are related by*

$$y^k \nabla^B_{\delta_k} = \boldsymbol{\nabla}_D. \tag{3.18}$$

Proof. We prove the result for the operators acting on a Finsler vector field. The generalization to general Finsler tensor fields is then straightforward. We have

$$\begin{aligned} y^k \nabla^B_{\delta_k} X^i &= y^k \left(\delta_k X^i + \bar{\partial}_j N^i_k X^j \right) && (3.19)\\ &= y^k \delta_k X^i + y^k \bar{\partial}_k N^i_j X^j && (3.20)\\ &= y^k \delta_k X^i + N^i_j X^j = \boldsymbol{\nabla}_D X^i, && (3.21) \end{aligned}$$

where we have first used torsion-freeness of the nonlinear connection to switch the indices j and k, and then Euler's theorem, using homogeneity of N^i_j. □

3.3.2 The Chern-Rund connection

Given a Finsler space with fundamental tensor g_{ij}, the Chern-Rund connection ∇^C is the unique linear connection on π^*TM that is horizontally torsion-free,

and horizontally metric compatible, i.e.

$$T_{\text{hor}} = 0, \qquad \nabla^C_{\delta_i} g_{jk} = 0, \tag{3.22}$$

where the horizontal torsion is defined using the canonical nonlinear connection associated with F. One can show that the Christoffel symbols of the Chern-Rund connection are given by

$${}^C\Gamma^k_{ij} = \tfrac{1}{2} g^{k\ell} \left(\delta_i g_{\ell j} + \delta_j g_{i\ell} - \delta_\ell g_{ij} \right), \qquad {}^C\bar{\Gamma}^k_{ij} = 0. \tag{3.23}$$

Note that the formula for the horizontal Christoffel symbols is formally identical to the formula for the Levi-Civita Christoffel symbols with partial derivatives replaced by horizontal derivatives. Indeed, the proof is also completely analogous to the proof of the formula for the Levi-Civita Christoffel symbols, so we will not repeat it here. From this expression together with the definition of the horizontal derivatives and the property (1.11) of the Cartan tensor, it is straightforward to show that, just as for the Berwald Christoffel symbols, the geodesic spray coefficients (1.16) can be expressed in terms of the Chern-Rund Christoffel symbols as

$$G^k = {}^C\Gamma^k_{ij} y^i y^j = {}^B\Gamma^k_{ij} y^i y^j. \tag{3.24}$$

In fact, for both the Chern-Rund and Berwald connection of a Finsler metric, we have the stronger result that

$${}^C\Gamma^i_{k\ell} y^k = {}^B\Gamma^i_{k\ell} y^k = N^i_\ell. \tag{3.25}$$

For the Berwald connection, this follows immediately from (3.15) since the Christoffel symbols are symmetric. The justification of the equality with the expression involving the Chern-Rund connection will be postponed to Corollary 4.2.3 (or see [56]). It follows from (3.25) that we can extend the result of Proposition 3.3.1 for the Berwald connection to the Chern-Rund connection as well.

Proposition 3.3.2. *The dynamical covariant derivative $\mathbf{\nabla}_D$, the Berwald connection ∇^B, and the Chern-Rund connection ∇^C, all corresponding to the same Finsler metric, are related by*

$$y^k \nabla^C_{\delta_k} = y^k \nabla^B_{\delta_k} = \mathbf{\nabla}_D. \tag{3.26}$$

Proof. This follows immediately from the definitions and (3.25). □

Both the Berwald and the Chern-Rund connection will be useful when formulating the field equations of Finsler gravity in Chapter 7. There are two other notable linear connections in Finsler geometry and although we will not use these in this dissertation, we will mention them for completeness in the following summary of the important linear connections on a Finsler space. From this summary, we see that in order to have better metric compatibility one must sacrifice some torsion-freeness, and vice versa:

- the **Berwald connection** is defined uniquely by:
  - $T_{\text{hor}}(\delta_i, \bar{\partial}_j) = T_{\text{ver}}(\delta_i, \bar{\partial}_j) = 0$. *('as torsion-free as possible')*
- the **Chern-Rund connection** is defined uniquely by:
  - $\nabla_{\delta_i} g_{jk} = 0$. *(horizontal metric-compatibility)*
  - $T_{\text{hor}} = 0$ *(horizontally torsion-free)*
- the **Hashiguchi connection** is defined uniquely by:
  - $\nabla_{\bar{\partial}_i} g_{jk} = 0$. *(vertical metric-compatibility)*
  - $T_{\text{ver}}(\delta_i, \bar{\partial}_j) = T_{\text{ver}}(\bar{\partial}_i, \bar{\partial}_j) = 0$. *(vertically 'as torsion-free as possible')*
- the (linear) **Cartan connection** is defined uniquely by:
  - $\nabla_{\delta_i} g_{jk} = \nabla_{\bar{\partial}_i} g_{jk} = 0$ *(full metric-compatibility)*
  - $T_{\text{hor}}(\delta_i, \delta_j) = T_{\text{ver}}(\bar{\partial}_i, \bar{\partial}_j) = 0$ *(minimal torsion-freeness contraint).*

3.3.3 Covariant derivatives along curves

Given a connection ∇ on π^*TM and a Finsler vector field $X \in \Gamma(\pi^*TM)$, the covariant derivative of X along a curve $\eta : I \to \mathcal{A}, t \mapsto \eta(t) = (\gamma(t), v(t))$ can be defined in the standard way as

$$\frac{D}{Dt} X^k \equiv \nabla_{\dot{\eta}} X^k = \nabla_{\dot{\gamma}^i \partial_i + \dot{v}^i \bar{\partial}_i} X^k = \dot{\gamma}^i \nabla_{\partial_i} X^k + \dot{v}^i \nabla_{\bar{\partial}_i} X^k, \tag{3.27}$$

where it is understood that the expressions are evaluated along the curve, i.e. at $(\gamma(t), v(t)) \in \mathcal{A} \subset TM$.

Now let us assume that the covariant derivative is given by either the Chern-Rund or the Berwald connection on a Finsler space. Expressing the partial

derivatives in terms of the horizontal and vertical derivatives, and using the definition (3.4) of the Christoffel symbols together with the fact that $\bar{\Gamma}^k_{ij} = 0$ for both connections, we obtain

$$\frac{D}{Dt}X^k = \dot{\gamma}^i \delta_i X^k(\gamma, v) + \dot{\gamma}^i \Gamma^k_{ij}(\gamma, v) X^j + \left(\dot{v}^i + N^i_j(\gamma, v)\dot{\gamma}^j\right) \bar{\partial}_j X^k(\gamma, v) \tag{3.28}$$

If we further assume that the curve η is the lift to TM of a curve in M (necessarily $\gamma : I \to M$), in the sense that $v = \dot{\gamma}$, then we find that

$$\begin{aligned} \frac{D}{Dt}X^k &= \dot{\gamma}^i \delta_i X^k(\gamma, \dot{\gamma}) + \dot{\gamma}^i \Gamma^k_{ij}(\gamma, \dot{\gamma}) X^j + \left(\ddot{\gamma}^i + N^i_j(\gamma, \dot{\gamma})\dot{\gamma}^j\right) \bar{\partial}_j X^k(\gamma, \dot{\gamma}) && (3.29) \\ &= \dot{\gamma}^i \delta_i X^k(\gamma, \dot{\gamma}) + N^k_j(\gamma, \dot{\gamma}) X^j + \left(\ddot{\gamma}^i + N^i_j(\gamma, \dot{\gamma})\dot{\gamma}^j\right) \bar{\partial}_j X^k(\gamma, \dot{\gamma}) && (3.30) \\ &= \boldsymbol{\nabla}_D X^k + \left(\ddot{\gamma}^i + N^i_j(\gamma, \dot{\gamma})\dot{\gamma}^j\right) \bar{\partial}_j X^k(\gamma, \dot{\gamma}) && (3.31) \end{aligned}$$

in terms of the dynamical covariant derivative $\boldsymbol{\nabla}_D$ introduced in Section 2.4. Here we have used that $y^i \Gamma^k_{ij} = N^k_j$ for the Chern-Rund as well as the Berwald connection, i.e. (3.25). Hence if γ is a geodesic, the last term vanishes by the geodesic equation, and we have $\frac{D}{Dt}X^k = \boldsymbol{\nabla}_D X^k$. In fact, the calculation done here for vector fields generalizes to any type of tensor field. Hence we obtain the following result, which generalizes Lemma 2.4.1.

Proposition 3.3.3. *The Berwald and Chern-Rund covariant derivatives along a geodesic γ are both given by the dynamical covariant derivative*

$$\frac{D}{Dt} = \boldsymbol{\nabla}_D. \tag{3.32}$$

We remark, just to be clear, that by the covariant derivative along γ we mean, strictly speaking, the covariant derivative along its lift η. We immediately have the following corollary, where covariant constancy refers to covariant constancy with respect to either (or equivalently, both) of the Chern-Rund and Berwald connections.

Corollary 3.3.4. *A tensor on the pullback bundle is covariantly constant along geodesics if and only if its dynamical covariant derivative vanishes.*

This result will lead to a geometric interpretation of Landsberg spaces in Chapter 4.

3.4 Geodesic deviation

Using the properties of the Berwald and Chern-Rund connection developed in Section 3.3 we are now in a position to formulate the geodesic deviation equation for geodesics of a Finsler metric, which plays an important role in the 'derivation of'—or at least the motivation for—the field equations of Finsler gravity that we will encounter in Chapter 7.

Consider a smooth 1-parameter family of geodesics along a curve in a Finsler space M, in the sense that there exist real numbers $a < b, c < d$ and a smooth map $H : (a, b) \times (c, d) \to M$, $(s, t) \mapsto H(s, t)$ such that for each fixed value of t the curve $\gamma_t : s \mapsto H(s, t)$ is a geodesic. Let $u = u(s, t)$ be the vector field on M that is tangent to the geodesics, i.e. $u = H_*(\partial_s) = (\partial H^i/\partial s)\partial_i$ is the push-forward of ∂_s, and let v be the deviation vector field $v = H_*(\partial_t) = (\partial H^i/\partial t)\partial_i$. Then it can be shown that the following differential equation is satisfied, known as the *geodesic deviation* equation

$$\frac{D^2 v^k}{Ds^2} + R^k{}_i(u)v^i = 0, \tag{3.33}$$

where we have suppressed the x-arguments for clarity (as we will continue to do below). Here the *geodesic deviation operator* $R^k{}_i$ is given by $R^k{}_i = R^k_{ij}y^j$ in terms of the nonlinear curvature R^k_{ij}, see (2.14), and D/Ds is the (Chern-Rund or equivalenly, Berwald) covariant derivative along the geodesic, introduced in Section 3.3.3. By application of Proposition 3.3.3 and Lemma 2.4.1, this covariant derivative may be written as

$$\frac{Dv^k}{Ds} = \boldsymbol{\nabla}_D v^k = \frac{\mathrm{d}v^k}{\mathrm{d}s} + N^k_j(u)v^j. \tag{3.34}$$

Hence an alternative but equivalent form of the geodesic deviation equation (3.33) is given by

$$\boldsymbol{\nabla}^2_D v^k + R^k{}_i v^i = 0, \tag{3.35}$$

with the understanding that the expression is evaluated at $y = u$. For a proof of the geodesic deviation equation, we refer e.g. to [58, §5.7] or [59, §6.1].

CHAPTER 4

Berwald, Landsberg and Unicorn Metrics

In this last chapter of the preliminaries we discuss two important classes of Finsler metrics: Berwald metrics and Landsberg metrics. Berwald metrics are central to Part II of this dissertation, where we investigate the pseudo-Riemannian metrizability of such metrics, and also to Chapter 8 and Chapter 9 in Part III, which deal with solutions in Finsler gravity of Berwald type. Landsberg metrics are only slightly more general than Berwald metrics, but the exact meaning of the word 'slightly' here has not completely been settled and is the domain of the so-called unicorn problem, which we briefly discuss as well. In Chapter 10 we will employ Landsberg metrics of unicorn type in the context of cosmology to model a Finslerian expanding universe.

4.1 Berwald metrics

A Finsler space is said to be of Berwald type, or simply Berwald, if the canonical nonlinear connection (2.23) reduces to a (necessarily smooth) linear connection on TM, or in other words, an affine connection on the base manifold M, meaning that the connection coefficients are of the form $N_i^k = \Gamma_{ij}^k(x)y^j$ for certain coefficients $\Gamma_{ij}^k(x)$. As already alluded to in Chapter 2 it can be inferred from the transformation behavior of N_j^i that the functions Γ_{jk}^i have the correct transformation behavior to be the Christoffel symbols of a (torsion-free) affine connection on M. We will refer to this affine connection as the associated affine connection,

or simply *the affine connection* on the Berwald space. Below we specialize some of the important results obtained in Chapter 2 to the setting of Berwald spaces.

First, the parallel transport (2.7) and autoparallel equations (2.8) reduce in this case to the familiar equations

$$\dot{V}^i + \Gamma^i_{jk}(\gamma)\dot{\gamma}^j V^k = 0, \qquad \ddot{\gamma}^i + \Gamma^i_{jk}(\gamma)\dot{\gamma}^j\dot{\gamma}^k = 0, \tag{4.1}$$

and indeed the relation between the affine Christoffel symbols and the spray coefficients is given by (cf. (3.24))

$$G^i = \Gamma^i_{jk}(x)y^j y^k. \tag{4.2}$$

The curvature tensors of a Berwald space can be written as

$$R^k_{ij} = \bar{R}^k{}_{\ell ij}(x)y^\ell, \quad \mathrm{Ric} = \bar{R}_{ij}(x)y^i y^j, \quad R_{ij} = \tfrac{1}{2}\left(\bar{R}_{ij}(x) + \bar{R}_{ji}(x)\right) \tag{4.3}$$

in terms of the curvature tensor and Ricci tensor, respectively, of the affine connection, defined by (2.17), i.e.

$$\bar{R}^k{}_{\ell ij} = \partial_i\Gamma^k_{j\ell} - \partial_j\Gamma^k_{i\ell} + \Gamma^k_{im}\Gamma^m_{j\ell} - \Gamma^k_{jm}\Gamma^m_{i\ell}, \qquad \bar{R}_{lk} = \bar{R}^i{}_{lik}. \tag{4.4}$$

We will also refer to $\bar{R}^k{}_{\ell ij}$ and $\bar{R}_{lk}$ as the *affine curvature tensor* and *affine Ricci tensor*, respectively. Since the last identity in (4.3) will prove to be of particular interest later, we restate it as a lemma.

Lemma 4.1.1. *The Finsler-Ricci tensor of a Berwald space is the symmetrization of the affine Ricci tensor.*

It is worth pointing out that in the classical positive definite case with $\mathcal{A} = TM_0$, it is not necessary to symmetrize since in that case the affine Ricci tensor is always symmetric. This is a consequence, for instance, of Szabo's metrization theorem. In other signatures or, more accurately, in cases where $\mathcal{A} \neq TM_0$, the symmetrization is really necessary. We will come back to these matters in detail in Chapter 6

Berwald manifolds can be characterized in many equivalent ways, some of which we state below. We will only sketch the proof of their equivalences. In Chapter 5 we will encounter an additional equivalent characterization of Berwald spaces, introduced in [H1], based on an auxiliary pseudo-Riemannian metric.

Theorem 4.1.2. *The following are all equivalent characterizations of Berwald manifolds:*

(i) The canonical nonlinear connection of F is linear;

(ii) *F admits a (necessarily unique) smooth, torsion-free, metric-compatible affine connection;*

(iii) *The Christoffel symbols of the Chern-Rund connection do not depend on y, i.e. ${}^{C}\Gamma^k_{ij} = {}^{C}\Gamma^k_{ij}(x)$;*

(iv) *The Christoffel symbols of the Berwald connection do not depend on y, i.e. ${}^{B}\Gamma^k_{ij} = {}^{B}\Gamma^k_{ij}(x)$;*

(v) *The spray is quadratic in y;*

(vi) $\bar{\partial}_j\bar{\partial}_k\bar{\partial}_\ell G^i = 0$.

Proof sketch. $(i) \Leftrightarrow (iv)$ follows immediately from the definition of the Berwald Christoffel symbols in terms of the nonlinear connection coefficients, and $(iv) \Leftrightarrow (v)$ follows immediately from (3.16) and the 2-homogeneity of G^k. $(v) \Leftrightarrow (vi)$ is obvious. $(i) \Leftrightarrow (ii)$ follows from the fact that the canonical connection is the unique torsion-free metric-compatible homogeneous connection on TM. The equivalence with (iii) requires some more work and for this, we refer to [56]. For more detailed proofs of the other equivalences we also refer to [57] and [41]. □

From the fact that the Berwald and Chern-Rund Christoffel symbols only depend on x for a Berwald metric, together with the fact that the spray can be expressed in the same way in terms of both, namely as $G^k = \Gamma^k_{ij}(x)y^iy^j$, cf. (3.17) and (3.24), and finally the fact that both connections are horizontally torsion-free, it follows that the Berwald and Chern-Rund Christoffel symbols coincide for Berwald metrics. Moreover, their horizontal Christoffel symbols coincide with the Christoffel symbols of the (unique) torsion-free, metric-compatible affine connection.

4.2 Landsberg metrics

We have seen above that for Berwald spaces, the difference between the Berwald and Chern-Rund connection vanishes. In general, this difference defines a tensor, called the *Landsberg tensor*. A Landsberg metric is defined by the property that this tensor vanishes. In other words, a Landsberg metric is one for which the Berwald and Chern-Rund connection coincide. But as for Berwald metrics, there are several equivalent and perhaps more enlightening characterizations of Landsberg spaces.

Definition and Theorem 4.2.1. *The Landsberg curvature S_{ijk} tensor is given by any and each of the following equivalent definitions:*

$$S_{jkl} = -\tfrac{1}{4} y_i \, \bar{\partial}_j \bar{\partial}_k \bar{\partial}_\ell G^i \qquad \text{(in terms of the spray)} \tag{4.5}$$

$$= g_{ij} \left({}^{B}\Gamma^i_{k\ell} - {}^{C}\Gamma^i_{k\ell} \right) \qquad \text{(Berwald minus Chern-Rund } \Gamma\text{-symbols)} \tag{4.6}$$

$$= \boldsymbol{\nabla}_D C_{jkl} \qquad \text{(dynamical cov. der. of the Cartan tensor)} \tag{4.7}$$

$$= -\tfrac{1}{2} \nabla^B_{\delta_j} g_{kl}. \qquad \text{(hor. Berwald cov. der. of } g_{ij}) \tag{4.8}$$

We do not attempt to give a complete proof of these equivalences here but refer instead to [56] for details. Clearly, it follows that the Landsberg tensor is completely symmetric. The *mean Landsberg curvature* is defined as the trace $S_j = g^{k\ell} S_{jk\ell}$ of the Landsberg tensor and can in particular be expressed as

$$S_j = \boldsymbol{\nabla}_D C_j = \boldsymbol{\nabla}_D \left(\bar{\partial}_j \ln \sqrt{|\det g|} \right). \tag{4.9}$$

With Definition and Theorem 4.2.1 we can define the notion of a *Landsberg space* as follows:

Definition 4.2.2. *A Landsberg space is a Finsler space that satisfies any and hence each of the following equivalent conditions:*

- $S_{jkl} = -\tfrac{1}{4} y_i \, \bar{\partial}_j \bar{\partial}_k \bar{\partial}_\ell G^i = 0$
- *The Berwald and the Chern-Rund Christoffel symbols coincide*
- $\boldsymbol{\nabla}_D C_{jkl} = 0$, *i.e. the Cartan tensor is covariantly constant along geodesics.*
- $\nabla^B_{\delta_k} g_{ij} = 0$, *i.e. the fundamental tensor is horizontally constant with respect to the Berwald connection.*

The interpretation in terms of covariant constancy (with respect to the Berwald as well as Chern-Rund connection) along geodesics in the third equivalent condition is justified by Corollary 3.3.4. As a corollary of Definition and Theorem 4.2.1 we also obtain the following result that we already used in Section 3.3 (see (3.25)), but the justification of which we had postponed.

Corollary 4.2.3. *Both the Chern-Rund and Berwald connection satisfy*

$${}^{C}\Gamma^i_{k\ell} y^k = {}^{B}\Gamma^i_{k\ell} y^k = N^i_\ell. \tag{4.10}$$

Proof. It follows from the Leibniz rule for $\boldsymbol{\nabla}_D$, the fact that $\boldsymbol{\nabla}_D y^k = 0$ and the fact that $y^k C_{ijk} = 0$ that $y^k \boldsymbol{\nabla}_D C_{ijk} = 0$. It thus follows from the theorem that $y^k g_{ij} \left({}^B\Gamma^i_{k\ell} - {}^C\Gamma^i_{k\ell} \right) = 0$ and it then follows by nondegeneracy of g_{ij} that ${}^C\Gamma^i_{k\ell} y^k = {}^B\Gamma^i_{k\ell} y^k$, the latter being equal to N^i_ℓ, by (3.15). □

Finally, we say that a Finsler space is *weakly Landsberg* if the mean Landsberg curvature vanishes. It thus follows immediately from the definitions that we have the following inclusions:

$$\text{pseudo-Riemannian} \subset \text{Berwald} \subset \text{Landsberg} \subset \text{Weakly Landsberg}.$$

4.3 The unicorn problem

It has been a long-standing open question whether the inclusion Berwald $\subset$ Landsberg is strict. Do there exist Landsberg spaces that are not Berwald? This question, along with some variations of it, is known as the *unicorn problem*. For the important case where the domain is all of $\mathcal{A} = TM_0$ the answer is unknown. If the condition $\mathcal{A} = TM_0$ is relaxed, however, some examples of non-Berwaldian Landsberg spaces have been obtained, but such examples are still exceedingly rare. As such, non-Berwaldian Landsberg spaces are referred to as *unicorns* [46]. We recommend [46, 70] for reviews on the unicorn problem.

The first unicorns were found by Asanov [47] in 2006 and his results were generalized by Shen [48] a few years later. These were the only known examples of unicorns until Elgendi very recently provided some additional examples [49]. One of the families of unicorns introduced by Elgendi will be our point of departure in Chapter 10. In fact, by modifying Elgendi's metrics slightly in such a way that they retain the unicorn property, we will find that there exist exact solutions to Pfeifer and Wohlfarth's field equation of Finsler gravity, of unicorn type, that can be interpreted in a cosmological setting.

Part II

Berwald Spaces and Pseudo-Riemann Metrizability

Introduction to Part II: Berwald Spaces and Pseudo-Riemann Metrizability

In Part I we have introduced the notion of a Berwald space and reviewed some of the standard characterizations of such spaces. Here, in Part II, we obtain a novel characterization of Berwald spaces and use it, among other things, to study the question of pseudo-Riemann metrizability, i.e. the question of whether the canonical torsion-free metric-compatible connection on a Berwald space can be understood as the Levi-Civita connection of some auxiliary pseudo-Riemannian metric. While a well-known theorem due to Szabo [42] states that this is always possible in the classical positive definite scenario, the question of metrizability in other signatures, such as Lorentzian signature, and in scenarios with less stringent smoothness constraints, is still almost completely unexplored.

Part II starts with Chapter 5, where we present a novel necessary and sufficient condition for a Finsler metric to be of Berwald type and we explore its consequences. This characterization makes use of an arbitrary auxiliary pseudo-Riemannian metric, which renders it especially useful in situations with a preferred pseudo-Riemannian metric. This is the case, for instance, when dealing with Finsler metrics that are (anisotropically) conformal to a pseudo-Riemannian metric, or when dealing with (α, β)-metrics. (α, β)-metrics are simply Finsler metrics constructed from a pseudo-Riemannian metric α and a 1-form β and such metrics will play a major role throughout both Part II and Part III. Indeed, we obtain as a corollary a very useful necessary and sufficient Berwald condition for (α, β)-metrics and we apply this result to several specific cases of interest: Randers metrics, exponential metrics, and generalized m-Kropina metrics, reproducing some well-known results as well as presenting new ones. Our

generalized m-Kropina metrics are a generalization of the standard m-Kropina metric; they naturally appear as the most general (α, β)-metrics consistent with a certain form of the affine connection.

Next, Chapter 6 is devoted to the pseudo-Riemann metrizability of Finsler metrics in arbitrary signatures. We show that Szabo's metrization theorem cannot be extended to arbitrary signatures such as Lorentzian signature. We then go on to obtain necessary and sufficient conditions for (generalized) m-Kropina spaces to be locally metrizable, resulting in a complete characterization of locally metrizable m-Kropina spaces, and a partial characterization of locally metrizable generalized m-Kropina spaces. We end the chapter with a classification of locally metrizable Ricci-flat m-Kropina spaces.

CHAPTER 5

Characterization of Berwald Spaces

The purpose of this chapter is to provide a novel characterization of Berwald spaces, Theorem 5.1.1, and explore its consequences, including new results as well as well-known ones that follow from it particularly easily or that are somehow relevant to later chapters. As an important application, a major part of the chapter will be devoted to (α, β)-metrics of Berwald type. Many of the results discussed here have been published in [H1], albeit in many cases in a substantially different form. In particular, the results are presented here with a heightened focus on mathematical rigor. Apart from the applications that we discuss here, Theorem 5.1.1 has also been used to classify all spatially homogeneous and isotropic Berwald spacetimes, i.e. Berwald spacetimes with cosmological symmetry [71, 72].

5.1 The Berwald condition

In [H1] we introduced a new necessary and sufficient condition for a Finsler metric to be of Berwald type, a condition that employs an auxiliary pseudo-Riemannian metric. Here we provide an alternative proof to that which is given in the original paper.

Theorem 5.1.1. *Let $(M, \mathcal{A}, L)$ be a Finsler space and let A be an auxiliary pseudo-Riemannian Finsler Lagrangian on M such that $L = \Omega A$ on $\mathcal{A}$ for some $\Omega \in C^\infty(\mathcal{A})$. Then:*

(i) L is of Berwald type if and only if there exists a smooth symmetric classical $(1,2)$-tensor field $T^j{}_{ik} = T^j{}_{ki}$ on M, such that

$$A\mathring{\delta}_i\Omega = T^j{}_{ik}y^k\bar{\partial}_j L, \tag{5.1}$$

where $\mathring{\delta}_i$ denotes the horizontal derivative with respect to the Levi-Civita connection of A.

(ii) If so, the Christoffel symbols of the torsion-free L-compatible affine connection are given by

$$\Gamma^j_{ik} = \mathring{\Gamma}^j{}_{ik} + T^j{}_{ik}, \tag{5.2}$$

where $\mathring{\Gamma}^j{}_{ik}$ are the Levi-Civita Christoffel symbols of A. Moreover the Christoffel symbols Γ^j_{ik} coincide with the horizontal Chern-Rund as well as Berwald Christoffel symbols.

Note that given a Finsler Lagrangian L, any positive definite A will do. In that case there always exists an Ω such that $L = \Omega A$, namely $\Omega = L/A$. However, as one might expect, the theorem is most useful in situations where there is some preferred pseudo-Riemannian metric A. Typical examples of this are Finsler metrics that are (anisotropically) conformal[1] to a pseudo-Riemannian metric, and (α,β)-metrics, which we study in detail in the next two sections. We also point out that whenever $A \neq 0$ the condition (5.1) can be written as

$$\partial_i\Omega - \mathring{\Gamma}^j_{ik}y^k\bar{\partial}_j\Omega = T^j{}_{ik}y^k\left(\bar{\partial}_j\Omega + \frac{2y_j\Omega}{A}\right), \tag{5.3}$$

which is how it originally appeared in [H1]. Here indices are lowered with the metric A.

Proof of Theorem 5.1.1. One of the equivalent characterizations of a Berwald metric (see Theorem 4.1.2) is that there exists a (necessarily unique) torsion-free linear connection on TM, with Christoffel symbols denoted by Γ^k_{ij}, compatible with L, i.e. such that L is horizontally constant with respect to it. Denoting the corresponding (linear) nonlinear connection by N^j_i, this means that

$$\delta_i L = \partial_i L - N^j_i\bar{\partial}_j L = \partial_i L - \Gamma^j_{ik}y^k\bar{\partial}_j L = 0. \tag{5.4}$$

[1]The adverb 'anisotropically' refers to the fact that the conformal factor is allowed to depend not only on x but also on y.

After substituting $L = \Omega A$ this can be written as

$$\left(\partial_i \Omega - \Gamma^j_{ik} y^k \bar{\partial}_j \Omega\right) A = -\Omega \left(\partial_i A - \Gamma^j_{ik} y^k \bar{\partial}_j A\right). \tag{5.5}$$

If $\mathring{\Gamma}^j_{ik}$ denotes the Levi-Civita Christoffel symbols of the metric A, then we can write $\Gamma^j_{ik} = \mathring{\Gamma}^j_{ik} + T^j_{ik}$, for some symmetric (in $i \leftrightarrow k$) classical tensor field T^j_{ik}, since the difference between two linear connections is necessarily tensorial. Since $\tilde{\Gamma}^j_{ik}$ is compatible with A it follows that the term in between brackets on the RHS of (5.5) reduces to

$$\partial_i A - \Gamma^j_{ik} y^k \bar{\partial}_j A = \underbrace{\partial_i A - \mathring{\Gamma}^j_{ik} y^k \bar{\partial}_j A}_{=0} - T^j_{ik} y^k \bar{\partial}_j A = -T^j_{ik} y^k \bar{\partial}_j A, \tag{5.6}$$

and we can rewrite (5.5) as

$$\left(\partial_i \Omega - \mathring{\Gamma}^j_{ik} y^k \bar{\partial}_j \Omega\right) A = T^j_{ik} y^k \left(A \bar{\partial}_j \Omega + \Omega \bar{\partial}_j A\right) = T^j_{ik} y^k \bar{\partial}_j L. \tag{5.7}$$

Recognizing on the LHS the $\mathring{\Gamma}$-horizontal derivative of Ω, the first result follows. Since $\Gamma^j_{ik} = \mathring{\Gamma}^j_{ik} + T^j_{ik}$, the formula for the Christoffel symbols of the L-compatible linear connection follows immediately as well, and as discussed in Section 4.1, these have to coincide with the Chern-Rund as well as Berwald horizontal Christoffel symbols. □

An immediate consequence of the theorem is a characterization of spray invariance under (anisotropically) conformal transformations of a pseudo-Riemannian metric, first obtained by Tavakol and Van den Bergh [15]. It follows essentially by setting $T^j_{ik} = 0$.

Corollary 5.1.2. *Suppose that $L = \Omega A$ is a Finsler Lagrangian on $\mathcal{A} = TM_0$, where A is a pseudo-Riemannian metric and $\Omega \in C^\infty(\mathcal{A})$. Then the geodesic spray of L coincides with that of A, i.e. $G^k = \mathring{G}^k = \mathring{\Gamma}^k_{ij} y^i y^j$, if and only if Ω is horizontally constant w.r.t. A, i.e.*

$$\mathring{\delta}_i \Omega = \partial_i \Omega - \mathring{\Gamma}^b{}_{ik} y^k \bar{\partial}_b \Omega = 0. \tag{5.8}$$

Proof. If the geodesic spray of L coincides with that of A then $T^j_{ik} = 0$ and it follows by the theorem that (5.8) must hold whenever $A \neq 0$. Since $TM \backslash \{A = 0\}$ is dense in $\mathcal{A} = TM_0$ it follows by continuity of $\mathring{\delta}_i \Omega$ that (5.8) must in fact hold everywhere on $\mathcal{A}$. Conversely, suppose that (5.8) holds. Then condition (5.1)

holds with $T^j{}_{ik} = 0$. Hence, by the theorem, L is Berwald and the spray of L coincides with that of A. □

5.2 (α, β)-metrics of Berwald type

5.2.1 (α, β)-metrics

One instance in which Theorem 5.1.1 is particularly useful is when dealing with (α, β)-metrics. Here $\alpha = \sqrt{|a_{ij}(x)y^i y^j|}$ and $\beta = b_i(x)y^i$ are scalar variables on TM defined in terms of a (smooth) pseudo-Riemannian metric $a = a_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ on M and a (smooth) 1-form $b = b_\mu \mathrm{d}x^\mu$ on M, and an (α, β)-metric is essentially a Finsler metric that is constructed only from some α and β, i.e. $F = f(\alpha, \beta)$ for some function f, or more precisely, $F(x, y) = f(\alpha(x, y), \beta(x, y))$. Due to homogeneity it follows that whenever $\alpha \neq 0$ on $\mathcal{A}$ any such F can be written in the standard form $F = \alpha\phi(\beta/\alpha)$ for a smooth function ϕ. In what follows we will always tacitly assume that $\alpha \neq 0$ on $\mathcal{A}$, which is the case for many examples of interest. We will therefore work with the following definition.

Definition 5.2.1. *An (α, β)-metric is a Finsler metric with domain $\mathcal{A}$ not including $\alpha = 0$ that can be written in the form $F = \alpha\phi(\beta/\alpha)$, where ϕ is smooth on its domain.*

We will often denote β/α by s, so that we can view ϕ as a function of the real variable s.

In Appendix C we derive an expression (10.1.4) for the fundamental tensor of a not necessarily positive definite (α, β)-metric and its determinant. From this, it follows in particular that ϕ never vanishes. In accordance with the notation used above, we will denote by $A = a_{ij}y^i y^j$ the Finsler Lagrangian corresponding to α, or rather to a_{ij}. By some abuse of language, we will sometimes refer to either of α, A, a and a_{ij} as the pseudo-Riemannian metric, and to β, b and b_i as the 1-form, and sometimes we will even write expressions such as $\beta = \mathrm{d}t$ and

$$\alpha = \sqrt{|-\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2|}, \tag{5.9}$$

but the interpretation of such expressions should be clear. We also denote

$$|b|^2 = a_{ij}b^i b^j, \qquad \mathrm{sgn}(A) = \epsilon. \tag{5.10}$$

In the presence of an (α, β)-metric, indices will be raised and lowered with a_{ij} unless otherwise specified. Well-known examples of (α, β)-metrics are:

- pseudo-Riemannian Finsler metrics $F = \alpha$;
- Randers metrics $F = \alpha + \beta$;
- Kropina metrics $F = \frac{\alpha^2}{\beta}$;
- m-Kropina metrics $F = \alpha^{1+m}\beta^{-m}$, where $m \in \mathbb{R}$, also known as generalized Kropina metrics, Bogoslovsky metrics or Bogoslovsky-Kropina metrics.

From a physics perspective, (α, β)-metrics are very relevant as they provide a means of deforming pseudo-Riemannian metrics α into properly Finslerian Finsler metrics. And it turns out, as we will see in Part III, that these types of metrics can in fact be used to generalize some of the solutions to Einstein's field equations to properly Finslerian solutions to the field equation in Finsler gravity.

As a quick first application of Theorem 5.1.1 we obtain the following well-known result (see e.g. [73]), where we denote by $\mathring{\nabla}$ the Levi-Civita connection of α.

Corollary 5.2.2. *Let $F = \alpha\phi(\beta/\alpha)$ and suppose that $\mathring{\nabla}_i b_j = 0$. Then F is Berwald and the affine connection and geodesic spray of F coincide with those of α.*

Proof. It suffices to show that $\mathring{\delta}_i\Omega = 0$, for in that case condition (5.1) holds with ${T^j}_{ik} = 0$, showing F is Berwald with affine connection (and hence geodesic spray) equal to that of α. We have $\mathring{\delta}_i\Omega = \frac{\partial\Omega}{\partial\alpha}\mathring{\delta}_i\alpha + \frac{\partial\Omega}{\partial\beta}\mathring{\delta}_i\beta$, where $\mathring{\delta}_i\alpha = 0$ because A is horizontally constant w.r.t. its own Levi-Civita connection, by definition. And for any 1-form on M and any (linear) connection, it can easily be checked that $\mathring{\delta}_i\beta = (\mathring{\nabla}_i b_j)y^j$, and hence in the particular case at hand we also have $\mathring{\delta}_i\beta = 0$, resulting in $\mathring{\delta}_i\Omega = 0$ and completing the proof. □

5.2.2 The Berwald condition for (α, β)-metrics

The general Berwald condition for (α, β)-metrics can be obtained by applying (5.1) directly to Finsler metrics of the form $F = \alpha\phi(\beta/\alpha)$. We remind the reader that $\epsilon = \mathrm{sgn}(A)$ and that indices are raised and lowered with a_{ij}. Hence in what follows, $y_i = a_{ij}y^j$.

Theorem 5.2.3. *An (α, β)-metric $F = \alpha\phi(\beta/\alpha)$ is of Berwald type if and only if there exists a smooth tensor field $T^j{}_{ik}(x) = T^j{}_{ki}(x)$ on M such that*

$$y^j \mathring{\nabla}_i b_j = T^j{}_{ik} y^k \left[b_j + \frac{\epsilon}{\alpha} \left(\frac{\phi}{\phi'} - \frac{\beta}{\alpha} \right) y_j \right] \tag{5.11}$$

whenever $\phi' \neq 0$ and $T^j{}_{ik} y^k y_j = 0$ whenever $\phi' = 0$. In that case, the affine connection is given by

$$\Gamma^j_{ik} = \mathring{\Gamma}^j{}_{ik} + T^j{}_{ik}, \tag{5.12}$$

where $\mathring{\Gamma}^j{}_{ik}$ are the Christoffel symbols of α.

Proof. The Lagrangian corresponding to F reads $L = F^2 = \epsilon A \phi(\beta/\alpha)^2$, so we identify $\Omega = \epsilon\phi^2$, which is smooth on $\mathcal{A}$ since this set does not include $A = 0$, by assumption. We start by noting that, since $\mathring{\delta}_i \alpha = 0$, we have

$$\mathring{\delta}_i \Omega = \Omega' \mathring{\delta}_i \left(\frac{\beta}{\alpha} \right) = \frac{\Omega'}{\alpha} \mathring{\delta}_i \beta = \frac{\Omega'}{\alpha} (\mathring{\nabla}_i b_j) y^j, \tag{5.13}$$

where $\Omega' = \mathrm{d}\Omega/\mathrm{d}s$. On the other hand, making use of the identities

$$\bar{\partial}_i \alpha = \frac{\epsilon y_i}{\alpha}, \qquad \bar{\partial}_i s = \frac{1}{\alpha} \left(b_i - \frac{\beta y_i}{A} \right), \tag{5.14}$$

a straightforward computation shows that

$$\bar{\partial}_j L = \Omega' \alpha \left(b_j - \epsilon \frac{\beta}{\alpha^2} y_j \right) + 2\epsilon\Omega y_j \tag{5.15}$$

$$= \Omega' \alpha b_j + \epsilon \left(2\Omega - \frac{\beta}{\alpha} \Omega' \right) y_j. \tag{5.16}$$

Plugging (5.13) and (5.16) into (5.1) yields

$$\alpha\Omega' \left(\mathring{\nabla}_i b_j \right) y^j = T^j{}_{ik} y^k \left[\Omega' \alpha b_j + \epsilon \left(2\Omega - \frac{\beta}{\alpha} \Omega' \right) y_j \right]. \tag{5.17}$$

It follows from nondegeneracy and (C.14) that $\phi \neq 0$, and since we are also assuming that $\alpha \neq 0$, we can divide by $2\epsilon\alpha\phi$. Using $\Omega = \epsilon\phi^2$ and $\Omega' = 2\epsilon\phi\phi'$ the equation can thus be rewritten as

$$\phi' y^j \mathring{\nabla}_i b_j = T^j{}_{ik} y^k \left[\phi' b_j + \frac{\epsilon}{\alpha} \left(\phi - \frac{\beta}{\alpha} \phi' \right) y_j \right] \tag{5.18}$$

and by Theorem 5.1.1, F is Berwald iff (5.18) is satisfied for all $(x, y) \in \mathcal{A}$. Now we distinguish the two cases. If (x, y) is such that $\phi' \neq 0$ then we can divide by ϕ', leading directly to (5.11). If, on the other hand, $\phi' = 0$ then the condition (5.18) reduces to $T^j{}_{ik} y^k \frac{\epsilon}{\alpha} \phi y_j = 0$, or equivalently, $T^j{}_{ik} y^k y_j = T_{jik} y^k y^j = 0$. Note that since the set $\{\phi' = 0\}$ is not necessarily open this does not imply that $T^j{}_{ik} = 0$. The formula for the Christoffel symbols follows from (5.2). □

This result is closely related to [73, Prop. 6.3.1.1]. In many cases of interest, the set $\{\phi' = 0\} \subset \mathcal{A}$ has empty interior or equivalently, the set $\tilde{\mathcal{A}} \equiv \mathcal{A} \setminus \{\phi' = 0\}$ is dense in $\mathcal{A}$. In such a scenario it suffices to consider only the situation for points satisfying $\phi' \neq 0$.

Theorem 5.2.4. *Let $F = \alpha\phi(\beta/\alpha)$ be an (α, β)-metric such that the set $\{\phi' = 0\} \subset \mathcal{A}$ has empty interior. Then TFAE:*

(i) F is of Berwald type;

(ii) there exists a smooth tensor field $T^j{}_{ik}(x) = T^j{}_{ki}(x)$ on M such that (5.11) holds whenever $\phi' \neq 0$.

In that case, the affine connection is given by (5.12).

Proof. The implication $(i) \Rightarrow (ii)$ follows immediately from Theorem 5.2.3. For the other implication, we start by noting that, since the subset $\phi' = 0$ of $\mathcal{A}$ has empty interior, the set $\tilde{\mathcal{A}} \equiv \mathcal{A} \setminus \{\phi' = 0\}$ is dense in $\mathcal{A}$. If (5.11) is satisfied whenever $\phi' \neq 0$, then F is Berwald on $\tilde{\mathcal{A}}$, interpreted as a conic subbundle of the tangent bundle $T\pi(\tilde{\mathcal{A}})$ of $\pi(\tilde{\mathcal{A}}) \subset M$. (Note that $\tilde{\mathcal{A}}$ is a conic subbundle of TM iff $\pi(\tilde{\mathcal{A}}) = M$.) Indeed, $\pi(\tilde{\mathcal{A}}) \subset M$ is open since the canonical projection map π is an open map and $\tilde{\mathcal{A}}$ is open. Thus, $\pi(\tilde{\mathcal{A}})$ is itself a smooth manifold and it is easily verified that $\tilde{\mathcal{A}}$ satisfies all axioms for a conic subbundle of $T\pi(\tilde{\mathcal{A}})$. Moreover, it will be useful to note that $\pi(\tilde{\mathcal{A}}) \subset \pi(\mathcal{A}) = M$ is dense, as the continuous image of a dense set.

As we observed above, F is Berwald on $\tilde{\mathcal{A}}$. Then $\bar{\partial}_j \bar{\partial}_k \bar{\partial}_\ell G^i = 0$ holds on $\tilde{\mathcal{A}}$, by Theorem 4.1.2 and so by continuity of the LHS, we must in fact have $\bar{\partial}_j \bar{\partial}_k \bar{\partial}_\ell G^i = 0$ on the closure of $\tilde{\mathcal{A}}$ (in $\mathcal{A}$), which is just $\mathcal{A}$. But then, by Theorem 4.1.2, F is Berwald on $\mathcal{A}$. This proves (ii).

We now prove the remaining statement in the theorem. F is Berwald on $\mathcal{A}$, so by (the proof of) Theorem 5.2.3, there exists a symmetric smooth (classical) tensor field $\tilde{T}^j{}_{ik}$ on M that satisfies (5.18) everywhere on $\mathcal{A}$. Moreover, the corresponding affine connection on M is given by $\Gamma^j_{ik} = \mathring{\Gamma}^j{}_{ik} + \tilde{T}^j{}_{ik}$. But we also know that the affine connection on $\pi(\tilde{\mathcal{A}}) \subset M$ is given by $\Gamma^j_{ik} = \mathring{\Gamma}^j{}_{ik} + T^j{}_{ik}$, so by

the uniqueness of the affine connection, we must have $\tilde{T}^j{}_{ik}(x) = T^j{}_{ik}(x)$ for all $x \in \pi(\tilde{\mathcal{A}})$. Thus $T^j{}_{ik}$ and $\tilde{T}^j{}_{ik}$ are continuous maps on M that coincide on the dense subset $\pi(\tilde{\mathcal{A}}) \subset M$ and as such they must coincide on M. It follows that the affine connection on all of M is given by $\Gamma^j_{ik} = \mathring{\Gamma}^j{}_{ik} + T^j{}_{ik}$, as desired. □

We can now extend the result of Corollary 5.2.2 in the opposite direction, reproducing the following well-known result.

Corollary 5.2.5. *A properly Finslerian (α, β)-metric has the same spray as α if and only if $\mathring{\nabla}_i b_j = 0$.*

Proof. One direction is given by Corollary 5.2.2. The other direction follows immediately by setting $T = 0$ in the Berwald condition (5.11) and performing a y-derivative, which is possible since the set $\{\phi' \neq 0\} \subset \mathcal{A}$ is open and also nonempty, since the latter would imply that F is pseudo-Riemannian. □

5.3 Important classes of (α, β)-metrics

5.3.1 Essential properties of α, α^2 and β

Our next goal is to apply Theorem 5.2.3 to some specific (α, β)-metrics, both obtaining new results and reproducing some standard ones. In order to do so, we require several lemmas that concern the irrationality of the function $\alpha = \sqrt{|A|}$ and the irreducibly and divisibility of the polynomials A and β, viewed as functions of y. We say that a polynomial is irreducible (or prime) if it cannot be written as a product of two polynomials both with degree > 0. We will use, without proof, the well-known result (see e.g. [74]) that the ring of multivariate polynomials over any field, and in particular over $\mathbb{R}$, is a factorial ring, i.e. a unique factorization domain (UFD), meaning that any such polynomial can be expressed uniquely (up to some constants) as a product of irreducible polynomials. In this sense, the irreducible polynomials play a role analogous to that of the prime numbers in number theory. In particular, this implies that if a product PQ of polynomials P, Q is divisible by an irreducible polynomial R then either P or Q must be divisible by R.

Lemma 5.3.1. *If $a = a_{ij}\mathrm{d}x^i\mathrm{d}x^j$ is a pseudo-Riemannian metric on M with $\dim M > 2$ then for any fixed $x \in M$, $A = a_{ij}y^i y^j$ is an irreducible polynomial in y.*

Proof. Suppose that A were not irreducible, i.e. that $A = PQ$, with P, Q two polynomials of deg > 0. Since $\deg A = 2$ this implies that $\deg P = \deg Q = 1$ and we may write $P = p_i y^i$ and $Q = q_i y^i$, since any constant terms in P, Q would render the product PQ inhomogeneous, whereas A is homogeneous of degree 2. Then differentiating $A = PQ$ twice yields $2a_{ij} = p_i q_j + q_i p_j$. This shows that a_{ij} has at most rank 2, because the image of any vector v^i will always be of the form $a_{ij}v^j = c_1 p_i + c_2 q_i$ with $c_1, c_2 \in \mathbb{R}$. Since we're assuming that $\dim M > 2$ this implies that a is degenerate and hence not a pseudo-Riemannian metric, which is a contradiction. □

Note that β is irreducible as well since any linear polynomial is.

Lemma 5.3.2. *Let $a = a_{ij}\mathrm{d}x^i\mathrm{d}x^j$ be a pseudo-Riemannian metric on M with $n = \dim M > 1$ and fix $x \in M$. Then the restriction of $\alpha = \sqrt{|a_{ij}(x)y^i y^j|}$ to any open subset in y-space $\mathbb{R}^n$ is an irrational function of y.*

Proof. Fix some x and suppose that there were some open subset U of y-space on which α were rational in y, i.e. $\alpha = P/Q$, with P, Q polynomials in y, where we will assume without of generality (WLOG) that the fraction P/Q is in lowest form, i.e. P and Q have no common polynomial divisors of degree > 0. Then $\alpha^2 = P^2/Q^2$. Note that α^2 is polynomial on both of the open sets $U_\pm = U \cap \{\pm a_{ij}(x)y^i y^j > 0\}$, and that $U_+ \cup U_-$ is necessarily nonempty since otherwise U could not be open. Suppose then WLOG that U_+ is nonempty (otherwise switch the labels U_+ and U_-).

Since α^2 is polynomial on U_+, and since $\alpha^2 = P^2/Q^2$, it follows that P^2/Q^2 must be polynomial on U_+ and hence (as a rational function) it must be polynomial on $\mathbb{R}^n$. Since P and Q and hence P^2 and Q^2 have no common polynomial divisors of degree > 0 this implies that Q must be of degree 0, i.e. constant, and by a redefinition of P we may then WLOG assume that $Q = 1$. Then $\alpha^2 = P^2$, and similar to the proof of Lemma 5.3.1 we may write $P = p_i y^i$ and differentiate twice to reveal that $a_{ij} = p_i p_j$, which shows that a_{ij} has at most rank 1. Since we're assuming that $\dim M > 1$ this implies that a is degenerate and hence not a pseudo-Riemannian metric, which is a contradiction. □

The phrase 'restriction of α to any open subset' in Lemma 5.3.2 is significant, since on its whole domain, $\alpha = \sqrt{|a_{ij}(x)y^i y^j|}$ would trivially be irrational as a consequence of the absolute value signs below the square root even if a_{ij} were degenerate. For instance, $a_{ij} = c_i c_j$ would lead to $\alpha = |c_i y^i|$, which is irrational on $\mathbb{R}$ but rational on its restriction to the open subset $c_i y^i > 0$. Such cases are excluded by the lemma.

Lemma 5.3.3. *Fix $x \in M$ with $\dim M > 2$, let $c_1, c_2 \in \mathbb{R}$ and consider the polynomial $P = c_1 A + c_2 \beta^2$ in y, where β is not identically vanishing at x. If $c_1 \neq 0$ then P is not divisible by β and if $c_2 \neq 0$ then P is not divisible by A.*

Proof. Suppose first for contradiction that $c_2 \neq 0$ and P *is* divisible by A. Then $c_1 A + c_2 \beta^2 = c_3 A$ for some constant c_3 and hence $c_2 \beta^2 = (c_3 - c_1)A$, which implies, since A is not divisible by β by Lemma 5.3.1, that $c_3 = c_1$, which implies that $c_2 = 0$, contrary to our assumption. Next, suppose for contradiction that $c_1 \neq 0$ and P is divisible by β. Then $c_1 A + c_2 \beta^2 = \beta Q$ for some linear polynomial Q, which implies that $A = \frac{1}{c_1}(Q - c_2 \beta)\beta$ and in particular implies that A can be factorized into linear polynomials. This contradicts Lemma 5.3.1, hence completing the proof. □

5.3.2 Randers metrics and exponential metrics

We now turn to some specific (α, β)-metrics, for which we will deduce, among other things, the precise Berwald condition, starting with the Randers metric [75], for which the result is well-known (see e.g. [73]).

Proposition 5.3.4. *A Randers metric $F = \alpha + \beta$ with $\dim M > 1$ is Berwald if and only if $\mathring{\nabla}_i b_j = 0$. In that case, the affine connection of F coincides with the Levi-Civita connection of α.*

Proof. For Randers metrics we have $\phi(s) = 1 + s$, $\phi'(s) = 1$, so the Berwald condition can be written as

$$y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j = T^j{}_{ik} y^k \frac{\epsilon y_j}{\alpha}. \tag{5.19}$$

If $\mathring{\nabla}_i b_j = 0$ then the condition is clearly satisfied by $T^j{}_{ik} = 0$ and hence F is Berwald. On the other hand, suppose that F is Berwald. Then the LHS of the Berwald condition is linear and, in particular, is a rational function in y, so the RHS must be as well. However, as the product of an irrational function $1/\alpha$ (by Lemma 5.3.2) with a (quadratic and hence) rational function, the RHS is irrational unless $T^j_{ik} y^k y_j = T_{jik} y^k y^j = 0$, which implies, by taking two y-derivatives, that $T_{jik} + T_{kij} = 0$. (Recall that indices are raised and lowered by a_{ij}, so $T_{jik} = a_{jl} T^l{}_{ik}$ depends only on x.) Christoffel's trick then shows that

$$T_{jik} = \tfrac{1}{2}\left[(T_{jik} + T_{kij}) + (T_{ikj} + T_{jki}) - (T_{kji} + T_{ijk})\right] = 0, \tag{5.20}$$

Hence $y^j \mathring{\nabla}_i b_j = 0$ and hence $\mathring{\nabla}_i b_j = 0$. The last statement in the proposition follows immediately from the fact that $T^k_{ij} = 0$. □

As a second quick application of Theorem 5.2.4 we consider a kind of exponential (α, β)-metrics given by $\phi(s) = e^{cs^2/2}$ for some constant c. To the best of our knowledge, the following is a new result, first obtained in [H1].

Proposition 5.3.5. *If b_i is a nowhere vanishing 1-form on M and $c \neq 0$ then $F = \alpha e^{c(\beta/\alpha)^2/2}$ is of Berwald type if and only if $\mathring{\nabla}_i b_j = 0$.*

Proof. Since $\phi'(s) = cs\phi = 0$ iff $s = 0$ iff $\beta = 0$, which cannot hold on an open neighborhood of a point (x, y) unless $b_i(x) = 0$, the subset $\{\phi' = 0\} \subset \mathcal{A}$ has empty interior and we may apply Theorem 5.2.4. The Berwald condition reduces to

$$y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j = \epsilon T^j{}_{ik} y^k y_j \left(\frac{\alpha^2 - c\beta^2}{c\,\alpha^2 \beta} \right), \tag{5.21}$$

whenever $\beta \neq 0$. Now the implication to the left is obvious, for if $\mathring{\nabla}_i b_j = 0$ then the condition is satisfied with $T^j{}_{ik} = 0$. Conversely, supposing that F is Berwald, the linear LHS of the condition implies that the RHS must be linear as well, so that both α^2 and β must be divisors of the polynomial $T^j{}_{ik} y^k y_j \, (\alpha^2 - c\beta^2)$. It follows by Lemma 5.3.3, however, that neither of these is a divisor of $(\alpha^2 - c\beta^2)$. Hence both α^2 and β must be divisors of $T^j{}_{ik} y^k y_j$, but if $T^j{}_{ik} y^k y_j \neq 0$ this is impossible since the latter is only quadratic and not cubic or higher. We must therefore have $T^j{}_{ik} y^k y_j = 0$. Hence, again taking two derivatives and applying Christoffel's trick as in the proof of Proposition 5.3.4, we conclude that $T^j{}_{ik} = 0$. Then it follows by Corollary 5.2.5 that $\mathring{\nabla}_i b_j = 0$, as desired. □

And as a final quick application before we move on to the more intricate case of generalized m-Kropina metrics, we consider another type of exponential (α, β)-metric, given by $\phi(s) = e^{cs}$. We present it here for the first time.

Proposition 5.3.6. *If $c \neq 0$ then $F = \alpha e^{c\beta/\alpha}$ is of Berwald type if and only if $\mathring{\nabla}_i b_j = 0$.*

Proof. We start by noting that $\phi'(s) = c\phi \neq 0$, since otherwise the metric would be degenerate, by (C.14). The Berwald condition can be written as

$$(\epsilon c \alpha^2)(y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j) + c\beta T^j{}_{ik} y^k y_j = \alpha T^j{}_{ik} y^k y_j. \tag{5.22}$$

Again the implication to the left is obvious, so we focus on the implication to the right. For each x, the LHS is rational in y, whereas the RHS is irrational unless $T^j{}_{ik}y^k y_j = 0$, by Lemma 5.3.2. It follows that the equation can only be satisfied if $T^j{}_{ik}y^k y_j = 0$. As we have seen twice before now, in the proofs of Proposition 5.3.4 and Proposition 5.3.5, this implies that $T^j{}_{ik} = 0$ and hence that $\mathring{\nabla}_i b_j = 0$, by Corollary 5.2.5. □

5.3.3 Generalized m-Kropina metrics

Next, we consider a generalization of the m-Kropina metric that we will refer to as the *generalized* m-Kropina metric, given by

$$F = \alpha\phi(\beta/\alpha), \qquad \phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}, \qquad m, c, d = \text{const.} \tag{5.23}$$

Such metrics were first introduced in [H1] as the most general Berwald metrics compatible with a particular form of T^j_{ik}, as will be made precise in Theorem 5.3.12 below, and they reduce to the m-Kropina class whenever $\pm = +$, $c = 1$ and $d = 0$ and to the standard Kropina class [76] whenever, additionally, $m = 1$. For generalized m-Kropina metrics we will always assume $\mathcal{A}$ to be contained in the subset of TM characterized by $s > 0$, $c + ds^2 > 0$, in order for any real power of those expressions to be defined (but note that this is strictly speaking not required if m is, for instance, an integer). This implies, in particular, that we may assume without loss of generality that b_i is nowhere vanishing since this would imply that $s = 0$ at some point in $\mathcal{A}$. We will take this fact for granted in what follows. Our next main goal is to obtain the Berwald condition for generalized m-Kropina metrics. To this end, we start with a somewhat technical lemma, where we set $\tilde{\mathcal{A}} \equiv \mathcal{A} \setminus \{\phi' = 0\}$.

Lemma 5.3.7. *For a properly Finslerian generalized m-Kropina metric with* $\dim M > 1$ *the following are true:*

- $cm \neq 0$*;*
- $\{\phi' = 0\} \subset \mathcal{A}$ *has empty interior;*
- $\pi(\tilde{\mathcal{A}}) = M$*. In particular, $\tilde{\mathcal{A}}$ is a conic subbundle of TM.*

Proof. We have $\phi = \pm s^{-m}(c + ds^2)^{(m+1)/2}$ and $\phi' = (ds^2 - cm)s^{-1-m}(c + ds^2)^{(-1+m)/2}$, so $\phi' = 0$ iff $ds^2 = cm$. Note that $c \neq 0$ (otherwise $F = c_1\beta$ for

some constant c_1 and the fundamental tensor would be degenerate) and $m \neq 0$ (otherwise F would not be properly Finslerian). Hence $cm \neq 0$.

We now claim that the set $\{\phi' = 0\} \subset \mathcal{A}$ has empty interior. This is trivially true if $d = 0$, so we focus on the case $d \neq 0$. Suppose the claim were false. Then there would be an open $U \subset \mathcal{A}$ such that $s^2 = cm/d$ for all $(x, y) \in U$, or equivalently, $(b_i y^i)^2 = (cm/d)|a_{ij} y^i y^j|$ for all $(x, y) \in U$. By differentiating twice this would imply that $b_i b_j = \pm(cm/d) a_{ij}$, which is impossible, since the LHS is degenerate, since $\dim M > 1$, while the RHS is nondegenerate, thus proving the claim.

Finally, it remains to be shown that $\pi(\tilde{\mathcal{A}}) = M$. Suppose this were not true. Then there exists $x_0 \in M$ such that $\phi'(s(x_0, y)) = 0$ for all $(x_0, y) \in \mathcal{A}$, which would imply that $(b_i y^i)^2 = (cm/d)|a_{ij} y^i y^j|$ for all $(x_0, y) \in \mathcal{A}$. Since the set of y for which (x_0, y) lies in the open set $\mathcal{A}$ is open, we can perform the same differentiation argument as before, leading to the contradiction $b_i(x_0) b_j(x_0) = \pm(cm/d) a_{ij}(x_0)$. Since $\tilde{\mathcal{A}}$ is also open and conic, this completes the proof. □

Proposition 5.3.8. *A properly Finslerian generalized m-Kropina metric $F = \alpha\phi(\beta/\alpha)$, $\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}$ with $n = \dim M > 2$ is Berwald if and only if there exists a smooth vector field $f^i(x)$ on M such that*

$$\mathring{\nabla}_i b_j = cm(f^k b_k) a_{ij} + (c + \epsilon d|b|^2) f_i b_j - (cm - \epsilon d|b|^2) b_i f_j - \epsilon d (f^k b_k) b_i b_j. \tag{5.24}$$

In that case, the affine connection is given by

$$\Gamma^k_{ij} = \mathring{\Gamma}^k_{ij} + a^{k\ell} \left(f_i h_{j\ell} + f_j h_{i\ell} - f_\ell h_{ij} \right), \quad h_{ij} \coloneqq \epsilon d b_i b_j - cm a_{ij}. \tag{5.25}$$

Whenever $|b|^2 \neq 0$ and $c + \epsilon d|b|^2 \neq 0$, the vector field f^i can be expressed explicitly as

$$f_i = \frac{\partial_i \ln ||b|^2|}{2(c + \epsilon d|b|^2)} \tag{5.26}$$

and moreover, in this case, $\mathrm{d}f = 0$.

Proof. In view of Lemma 5.3.7 we may apply Theorem 5.2.4, which says that F is Berwald if and only if there exists a symmetric smooth tensor field $T^j{}_{ik}(x)$ on M such that the condition

$$y^j \mathring{\nabla}_i b_j = T^j{}_{ik} y^k \left(b_j + \frac{\epsilon y_j}{\alpha} \frac{c(1+m)s}{ds^2 - cm} \right) \tag{5.27}$$

is satsfied on $\tilde{\mathcal{A}}$, i.e. whenever $\phi' \neq 0$, or equivalently, $ds^2 - cm \neq 0$. So suppose that F is Berwald, i.e. the latter is true. Equation (5.27) is equivalent to

$$(ds^2 - cm)\left(y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j\right) = \frac{\epsilon y_j}{\alpha} T^j{}_{ik} y^k c(1+m)s, \tag{5.28}$$

and substituting $s = \beta/\alpha$ and noting that $\epsilon\alpha^2 = A$, this is, in turn, equivalent to

$$(\epsilon d\beta^2 - cmA)\left(y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j\right) = c(1+m) y_j T^j{}_{ik} y^k \beta. \tag{5.29}$$

Since the RHS is divisible by β, the LHS must be as well. However, since $cm \neq 0$, it follows from Lemma 5.3.3 that the first factor $(\epsilon d\beta^2 - cmA)$ of the LHS is not divisible by β, which is an irreducible polynomial. Hence the second factor $y^j \nabla_i b_j - T^j{}_{ik} y^k b_j$ must be divisible by β for any $x \in \pi(\tilde{\mathcal{A}})$. But $\pi(\tilde{\mathcal{A}}) = M$, by Lemma 5.3.7, so there exists a vector field ℓ^i on M such that

$$\left(y^j \mathring{\nabla}_i b_j - T^j{}_{ik} y^k b_j\right) = \ell_i \beta. \tag{5.30}$$

Note that there is *a priori* no reason why ℓ_i should be smooth. Indeed, at this stage, all we can infer is that the product $\ell_i \beta$ must be smooth. We will see below that ℓ_i does turn out to be smooth, though. The formula above shows on the one hand, by taking a derivative, that

$$\mathring{\nabla}_i b_j = T^k_{ij} b_k + \ell_i b_j, \tag{5.31}$$

and on the other hand, by plugging the relation back into (5.29), that

$$(\epsilon d\beta^2 - cmA)\ell_i = c(1+m) y_j T^j{}_{ik} y^k, \tag{5.32}$$

or equivalently,

$$T^j{}_{ik} y^k y_j \equiv T_{jik} y^k y^j = f_i H, \tag{5.33}$$

where $\ell_i = c(1+m) f_i$ and $H = \epsilon d\beta^2 - cmA$. We claim that this implies that f^k is smooth. To see this, note that for every $x_0 \in M$ there is a $y_0 \in \mathbb{R}^n$ such that $(x_0, y_0) \in \mathcal{A}$ and $H(x_0, y_0) \neq 0$. For otherwise we would have $H(x_0, y) = 0$ for all y in some open set, so that, by taking two derivatives, we would obtain $\epsilon d b_i b_j = cm a_{ij}$, which is a contradiction as the LHS is degenerate while the RHS is nondegenerate. Since H is continuous, it follows that there is an open neighborhood $U \subset \mathcal{A}$ of (x_0, y_0) such that $H \neq 0$ on U. Then it follows from (5.33) that $f_i = (T^j{}_{ik} y^k y_j)/H$, viewed as a function of x *and* y, is smooth on

U. In particular, f_i is smooth at (x_0, y_0), and since f_i only has a dependence on x, not y, this means that f_i is smooth at (x_0, y) for any y, and since x_0 was arbitrary, that just means that f_i is smooth as a function of x and y, and hence it is smooth as a function on M. Hence f^k is smooth, proving the claim. Furthermore, taking two derivatives of (5.33) shows that

$$\tfrac{1}{2}\left(T_{jik} + T_{kij}\right) = f_i h_{jk}, \qquad h_{ij} \coloneqq \epsilon d b_i b_j - cm a_{ij}, \tag{5.34}$$

and using Christoffel's trick, we find that

$$\begin{aligned} T_{jik} &= \tfrac{1}{2}\left[(T_{jik} + T_{kij}) + (T_{ikj} + T_{jki}) - (T_{kji} + T_{ijk})\right] && (5.35)\\ &= f_i h_{jk} + f_k h_{ij} - f_j h_{ik}. && (5.36) \end{aligned}$$

Plugging this into (5.31), we find that

$$\mathring{\nabla}_i b_j = cm(f^k b_k) a_{ij} + (c + \epsilon d|b|^2) f_i b_j - (cm - \epsilon d|b|^2) b_i f_j - \epsilon d (f^k b_k) b_i b_j. \tag{5.37}$$

Thus what we have shown so far is that if F is Berwald then this condition must be satisfied on $\tilde{\mathcal{A}}$ for some smooth vector field f^k on M. But since the equation does not depend on y anymore and since $\pi(\tilde{\mathcal{A}}) = M$, the condition must in fact hold for all $x \in M$. Conversely, if this condition is satisfied for some smooth vector field f^k on M and all $x \in M$ then it is straightforward (although a little tedious) to check that (5.27) is satisfied on $\tilde{\mathcal{A}}$. This proves the 'if and only if' statement of the theorem. The formula for the Christoffel symbols follows directly from (5.36) by raising the first index.
Next, in order to derive the explicit formula for f_i, we contract (5.24) with b^j, leading to

$$\tfrac{1}{2}\partial_i |b|^2 = (c + \epsilon d|b|^2)|b|^2 f_i, \tag{5.38}$$

which leads to the desired formula provided that $|b|^2 \neq 0$ and $c + \epsilon d|b|^2 \neq 0$. Finally, it is now easy to check that $\partial_i f_j = \partial_j f_i$, meaning that f is in this case necessarily closed, $\mathrm{d}f = 0$. □

In the special case where the 1-form is known to be closed the Berwald condition simplifies considerably.

Proposition 5.3.9. *A properly Finslerian generalized m-Kropina metric $F = \alpha\phi(\beta/\alpha)$, $\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}$ with closed 1-form β and with $n = \dim M > 2$ is Berwald if and only if there exists a smooth function $p : M \to \mathbb{R}$*

such that

$$\mathring{\nabla}_i b_j = p\left\{cm|b|^2 a_{ij} + \left[c(1-m) + \epsilon d|b|^2\right] b_i b_j\right\} \tag{5.39}$$

In that case, the affine connection is given by

$$\Gamma^k_{ij} = {}^{\alpha}\Gamma^k_{ij} + p\left[\epsilon d b_i b_j b^k - cm\left(b_i\delta^k_j + b_j\delta^k_i - b^k a_{ij}\right)\right]. \tag{5.40}$$

Conversely, if (5.39) holds then β is closed.

Proof. F is Berwald iff the Berwald condition (5.24) holds for some smooth f^k. The Berwald condition implies that $(\mathrm{d}b)(\partial_i, \partial_j) = \partial_i b_j - \partial_j b_i = \mathring{\nabla}_i b_j - \mathring{\nabla}_j b_i = c(1+m)(f_i b_j - f_j b_i)$, so since b_i is closed, this expression vanishes and hence $f_i b_j = f_j b_i$ must hold for all i, j, which is only possible if f_i is proportional to b_i in the sense that $f_k = p b_k$ for some function p on M (this can be checked easily at any given point in M by choosing coordinates in which b_i has only one nonvanishing component at that point). In this case (5.24) reduces to (5.39) and (5.25) reduces to (5.40). Moreover, since f_k is smooth, pb_k must be smooth, and this implies that p is smooth whenever $b_k \neq 0$, and since b_k is nowhere vanishing, this implies that p is smooth on M. The opposite holds trivially as well: (5.39) implies that $\partial_i b_j - \partial_j b_i = \mathring{\nabla}_i b_j - \mathring{\nabla}_j b_i = 0$ and hence β is closed. □

The Berwald condition for m-Kropina spaces follows immediately by setting $\pm = +, c = 1, d = 0$. Again, under the additional assumption that the 1-form is closed, this condition simplifies considerably.

Corollary 5.3.10. *Let $F = \alpha^{1+m}\beta^{-m}$ be a properly Finslerian m-Kropina metric on a manifold M with dimension greater than two.*

- *F is of Berwald type if and only if there exists a smooth vector field f^i on M satisfying*

$$\mathring{\nabla}_j b_i = m(f_k b^k)a_{ij} + b_i f_j - m f_i b_j. \tag{5.41}$$

  In this case, the affine connection is given by

$$\Gamma^\ell_{ij} = \mathring{\Gamma}^\ell_{ij} + m a^{\ell k}\left(a_{ij} f_k - a_{jk} f_i - a_{ki} f_j\right). \tag{5.42}$$

- *If β is closed, then F is Berwald if and only if there exists a smooth function $c : M \to \mathbb{R}$ such that*

$$\mathring{\nabla}_j b_i = c\left[m|b|^2 a_{ij} + (1-m) b_i b_j\right], \tag{5.43}$$

In this case, the affine connection is given by

$$\Gamma^\ell_{ij} = \mathring{\Gamma}^\ell_{ij} + mc\left(a_{ij}b^\ell - \delta^\ell_j b_i - \delta^\ell_i b_j\right), \tag{5.44}$$

- *Conversely, (5.43) implies that β must be closed.*

The condition (5.41) has also been obtained in [73] under the assumptions $|b|^2 \neq 0$ (essentially replacing our assumption that $\dim M > 2$) and $A > 0$. The less general condition (5.43) was obtained in [44], where it was believed to be the general Berwald condition due to some details that were overlooked. In our paper [H3] we identified the error in the proof in [44] and corrected it; we refer to [H3] for the details of that discussion. We remark that our c in (5.43) is related to $C(x)$ in [44] by $C(x) = (1+m)c/2$ and that our power m is related to the power n in [44] by $n = -2m/(1+m)$.

Our next aim is to prove Theorem 5.3.12, which may be viewed in some sense as a generalization of Corollary 5.2.5. It shows that generalized m-Kropina metrics are the most general (α, β)-metrics that are consistent with a certain form of the affine connection. In the version of the theorem we prove below, we will assume that $\phi' \neq 0$ on $\mathcal{A}$. However, this assumption is superfluous and the theorem can be proven also without it. Since the proof below is already quite lengthy, we ignore such generalizations for the moment and refer instead to Appendix D.2 for details. In what follows, we will denote by s both the function $s : (x, y) \mapsto s(x, y) = \beta(x, y)/\alpha(x, y)$ as well as its function values $s \in \mathbb{R}$. So we will write, for instance, that $s \in s(U)$ if the *value* s lies in the image of the set U under the *function* s. Moreover, if $A \subset X$ is a subset of a topological space X, we will denote by A^0 and $\overline{A}$ its (topological) interior and closure, respectively. We start with a technical but important lemma.

Lemma 5.3.11. *Suppose that b_i is nowhere vanishing. If $U \subset \mathcal{A}$ is open and connected then the image $s(U)$ is an interval with nonempty interior and hence it satisfies, in particular, $s(U) \subset \overline{s(U)^0 \setminus C}$ for any finite set C.*

Proof. Since U is connected and s continuous, $s(U)$ is a connected subset of $\mathbb{R}$ and hence it is an interval $s(U) = I \subset \mathbb{R}$. Furthermore, it is an interval containing more than one point. For suppose it contained only one point s_0. Then we would have $\beta^2 = s_0\alpha^2 = s_0|A|$ for all $(x, y) \in U$. Now fix some x_0 such that $(x_0, y) \in U$ and define $V_\pm = \{y \in \mathbb{R}^n : (x_0, y) \in U, \pm A(x_0, y) > 0\}$. Then at least one of V_+ and V_- is open and nonempty. Assume WLOG that V_+ is open and nonempty. Then $\beta(x_0, y)^2 = s_0 A(x_0, y)$ for all $y \in V_+$ and hence

this equation between polynomials must hold for all $y \in \mathbb{R}^n$. Since $0 \neq \mathcal{A}$, by assumption, $s_0 \neq 0$. But then $s_0^{-1}\beta(x_0, y)^2 = A(x_0, y)$ and so A is divisible by β, contradicting Lemma 5.3.1. Hence I is an interval that contains at least two points and as such, the interior I^0 of I must be a nonempty open interval. Removing a finite set of points C from I^0 thus yields a union $I^0 \setminus C = \cup_{j=1}^{k}(a_j, b_j)$ of open intervals with

$$a_1 < b_1 = a_2 < b_2 = a_2 < \cdots < b_{k-1} = a_k < b_k. \tag{5.45}$$

It is then clear that $\overline{I^0 \setminus C} = [a_1, b_k]$ and hence that $I \subset \overline{I^0 \setminus C}$, completing the proof. □

As a result, if some ODE for $\phi(s)$ is satisfied for all $s \in s(U) \setminus C$, where $U \subset \mathcal{A}$ is open and C a finite set, then the ODE may be solved (explicitly) on the open set $s(U)^0 \setminus C$ using standard methods, and this uniquely determines ϕ on all of $s(U)$ by continuous extension.

Theorem 5.3.12. *Let $F = \alpha\phi(\beta/\alpha)$ be a properly Finslerian (α, β)-metric with $\dim M > 2$. Suppose furthermore that $\mathcal{A}$ is connected and that $\phi' \neq 0$ on $\mathcal{A}$. Then TFAE:*

(i) F is Berwald and there exist nowhere vanishing $\lambda, \rho, \sigma \in C^\infty(M)$ such that

$$T^k{}_{ij} = \lambda b^k b_i b_j + \rho\left(b_i\delta^k_j + b_j\delta^k_i\right) + \sigma b^k a_{ij}; \tag{5.46}$$

(ii) β is closed and there are nonvanishing constants c, d, m and a nowhere vanishing $p \in C^\infty(M)$ such that

$$\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2} \tag{5.47}$$

$$\mathring{\nabla}_i b_j = p\left\{cm|b|^2 a_{ij} + \left[c(1-m) + \epsilon d|b|^2\right] b_i b_j\right\} \tag{5.48}$$

In that case $\rho = -\sigma = -cmp$ and $\lambda = \epsilon dp$.

Proof. Assuming (ii) then, according to Proposition 5.3.9, F is Berwald and the affine connection is given by (5.40) and from this we infer that T^k_{ij} is clearly of the form (5.46) with $\lambda = \epsilon dp$ and $\rho = -\sigma = -cmp$. Next, suppose that (i) holds. Then the Berwald condition (5.11) must be satisfied on all of $\mathcal{A}$ (since ϕ' never

vanishes) with T^k_{ij} given by (5.46), in which case this condition reduces to

$$y^j \mathring{\nabla}_i b_j = \lambda|b|^2\beta b_i + 2\rho\beta b_i + \sigma|b|^2 y_i + \frac{\epsilon\psi}{\alpha}\left(\lambda\beta^2 b_i + \rho(\beta y_i + A b_i) + \sigma\beta y_i\right), \tag{5.49}$$

where $\psi = \phi/\phi' - s$. We can write this also as

$$y^j \mathring{\nabla}_i b_j - \left(\lambda|b|^2\beta b_i + 2\rho\beta b_i + \sigma|b|^2 y_i\right) = \frac{\epsilon\psi}{\alpha}\left(\lambda\beta^2 + \rho A\right) b_i + \frac{\epsilon\psi}{\alpha}(\rho+\sigma)\,\beta y_i, \tag{5.50}$$

where we have collected all manifestly linear terms on the LHS and grouped the RHS by a b_i term and a y_i term. Since the LHS is linear, the RHS must be so as well, and in fact, by Lemma B.0.2 this can only be achieved if both terms on the RHS are linear, i.e. if $\frac{\epsilon\psi}{\alpha}(\rho+\sigma)\,\beta$ is independent of y and if $\frac{\epsilon\psi}{\alpha}(\lambda\beta^2 + \rho A)$ is linear. Since $A = 0$ is excluded, by definition, from $\mathcal{A}$ and since $\mathcal{A}$ has only a single connected component, by assumption, $\mathcal{A}$ must be contained in one of the open sets $A > 0$ or $A < 0$. And that, in particular, implies that $\epsilon = \operatorname{sgn}(A)$ is just a constant on $\mathcal{A}$. Thus we need $\frac{\psi}{\alpha}(\rho+\sigma)\,\beta$ to be independent of y and $\frac{\psi}{\alpha}(\lambda\beta^2 + \rho A)$ to be linear.

Step 1) Claim: $\rho = -\sigma$

First we consider the y_i term, which is linear in y iff $\psi(\rho+\sigma)\,\beta/\alpha$ is independent of y. The only way this can be true, assuming F is properly Finslerian, is if $\rho = -\sigma$. To see this, suppose that $\rho \neq -\sigma$. Then we must have $\psi\beta/\alpha = \psi s =: \ell$ where ℓ is a priori a function of x, but since the LHS depends only on s, ℓ must actually be constant. Hence $\psi = \phi/\phi' - s = \ell/s$ for all $(x, y) \in \mathcal{A}$ with $s \neq 0$. As a differential equation in the variable s this must hold for all $s \in s(\mathcal{A})\setminus\{0\}$. Since $\mathcal{A}$ is open and connected, we can apply Lemma 5.3.11 and its consequences: we can integrate the differential equation for $\phi(s)$ on the interior of $s(\mathcal{A}) \setminus \{0\}$ and then extend the solution by continuity to all of $s(\mathcal{A})$, i.e. for all relevant values of s.

We will now apply this scheme. Assuming WLOG that $\ell \neq 0$ (for otherwise $\phi - s\phi' = 0$ and $\det g_{ij} = 0$, by (C.14)) we integrate to find $\ln|\phi| = \int (s + \ell/s)^{-1}\,\mathrm{d}s = \ln(|c(\ell + s^2)|)/2$, where c is an integration constant. Then $\phi = \sqrt{c(\ell + s^2)}$ for s in the interior of $s(\mathcal{A}) \setminus \{0\}$, and by the argument given above, ϕ must in fact have this form for all relevant values of s. Hence $F^2 = \alpha^2\phi^2 = c(\ell\alpha^2 + \beta^2)$ is pseudo-Riemannian. This is excluded in the premise of the theorem and hence it is a contradiction, so we must indeed have $\rho = -\sigma$.

Step 2) Deriving an ODE in s

Next, we consider the b_i term, setting $\sigma = -\rho$. This term is linear iff the coefficient of b_i is linear, which is the case iff

$$\bar{\partial}_i \bar{\partial}_j \left[\frac{\psi}{\alpha} \left(\lambda\beta^2 + \rho A \right) \right] = 0. \tag{5.51}$$

For the first derivative, we have

$$\begin{aligned}
\bar{\partial}_i \left[\frac{\psi}{\alpha} \left(\lambda\beta^2 + \rho A \right) \right] &= \frac{\psi' \bar{\partial}_i s}{\alpha} \left(\lambda\beta^2 + \rho A \right) - \frac{\psi}{\alpha^2} \bar{\partial}_i \alpha \left(\lambda\beta^2 + \rho A \right) && (5.52) \\
&\quad + \frac{\psi}{\alpha} \left(2\lambda\beta b_i + 2\rho y_i \right) && (5.53) \\
&= \frac{\psi'}{\alpha^2} \left(b_i - \frac{\beta y_i}{A} \right) \left(\lambda\beta^2 + \rho A \right) - \frac{\epsilon \psi y_i}{\alpha^3} \left(\lambda\beta^2 + \rho A \right) && (5.54) \\
&\quad + \frac{\psi}{\alpha} \left(2\lambda\beta b_i + 2\rho y_i \right) && (5.55) \\
&= \frac{1}{\alpha^2} \left[- \left(\frac{\beta\psi'}{A} + \frac{\epsilon\psi}{\alpha} \right) \left(\lambda\beta^2 + \rho A \right) + 2\alpha\rho\psi \right] y_i && (5.56) \\
&\quad + \{ b_i \text{ terms} \}, && (5.57)
\end{aligned}$$

where we have used the identities

$$\bar{\partial}_i \alpha = \frac{\epsilon y_i}{\alpha}, \qquad \bar{\partial}_i s = \frac{1}{\alpha} \left(b_i - \frac{\beta y_i}{A} \right), \tag{5.58}$$

and separated the y_i and the b_i parts, the latter of which are irrelevant. The reason for this is the following. By inspecting the situation (or working it out exactly) it is clear that the second derivative in (5.51) will be of the form $f a_{ij} + g h_{ij}$, where f, g are functions and h_{ij} is a linear combination of $b_i b_j, b_i y_j, y_i b_j$ and $y_i y_j$. Hence h_{ij} has at most rank 2, which is strictly less than the rank of a_{ij}. That means that (5.51), i.e. the equation $f a_{ij} + g h_{ij} = 0$, can only hold if $f = 0$, for otherwise we could write $a_{ij} = (g/f) h_{ij}$ and a_{ij} and h_{ij} would have the same rank. It turns out that it suffices for what we aim to prove to consider this coefficient f, as we will see shortly. Differentiation of the b_i terms in (5.57), however, will never yield terms proportional to a_{ij}. In fact, the only way we get terms proportional to a_{ij} is by directly differentiating y_i in (5.57). No other terms resulting from the product rule are proportional to a_{ij} either. This simplifies this calculation enormously because it means that in order to get the a_{ij}-coefficient f from (5.57) we effectively need only replace y_i by its derivative

$\bar{\partial}_j y_i = a_{ij}$ to obtain

$$\bar{\partial}_i \bar{\partial}_j \left[\frac{\psi}{\alpha} \left(\lambda \beta^2 + \rho A \right) \right] = \tag{5.59}$$

$$\frac{1}{\alpha^2} \left[- \left(\frac{\beta \psi'}{A} + \frac{\epsilon \psi}{\alpha} \right) \left(\lambda \beta^2 + \rho A \right) + 2 \alpha \rho \psi \right] a_{ij} + g h_{ij}. \tag{5.60}$$

As we argued above, the vanishing of this implies that we must have

$$0 = \frac{1}{\alpha} \left[- \left(\frac{\beta \psi'}{A} + \frac{\epsilon \psi}{\alpha} \right) \left(\lambda \beta^2 + \rho A \right) + 2 \alpha \rho \psi \right] \tag{5.61}$$

$$= - \left(\frac{s \psi'}{\epsilon \alpha^2} + \frac{\epsilon \psi}{\alpha^2} \right) \left(\lambda \beta^2 + \rho A \right) + 2 \rho \psi \tag{5.62}$$

$$= - \epsilon \left(s \psi' + \psi \right) \left(\lambda s^2 + \rho \epsilon \right) + 2 \rho \psi, \tag{5.63}$$

or, after rewriting,

$$\frac{\psi'}{\psi} = \frac{\eta - s^2}{s(s^2 + \eta)}, \qquad \eta \equiv \epsilon \rho / \lambda, \tag{5.64}$$

whenever $s \neq 0$ and $s^2 + \eta \neq 0$. Note that ψ never vanishes for otherwise $\phi - s\phi' = 0$ and $\det g_{ij} = 0$, by (C.14). As before, since the LHS depends only on s, the RHS should also not depend on x explicitly, which implies that η, which is a priori a function on M, should actually be a constant[2]. Hence (5.64) is an ODE for $\psi(s)$ that should be satisfied for all for all $s \in K \equiv s(\mathcal{A}) \setminus \{\sqrt{-\eta}, 0, -\sqrt{-\eta}\}$.

Step 3) Solving the ODE to obtain $\phi(s)$
In analogy with how we solved the differential equation for ϕ in step 1 of this proof, this one is solved uniquely, for all $s \in K^0$, by

$$\psi = \frac{\tilde{c} s}{s^2 + \eta}, \tag{5.65}$$

where $\tilde{c}$ is an integration constant. We find ϕ on K^0 via $\psi = \phi / \phi' - s$. This can

[2]We can invert (5.64) to obtain η as a function of s, except for values of s that satisfy $\psi + s\psi' = 0$, which can only constitute a set with empty interior, for otherwise F would be pseudo-Riemannian. Hence, by continuous extension, the formula for $\eta(s)$ must also hold on this set.

now be written as

$$\frac{\phi'}{\phi} = \left(\frac{\tilde{c}s}{s^2+\eta} + s\right)^{-1} = \frac{ds^2 - cm}{ds^3 + cs}, \tag{5.66}$$

where we have introduced an arbitrary constant d and defined $c = d(\tilde{c} + \eta)$ and $m = -d\eta/c = -\eta/(\eta + \tilde{c})$. This is uniquely solved by

$$\phi(s) = \tilde{d}s^{-m}(c + ds^2)^{(m+1)/2}, \tag{5.67}$$

where we may absorb the integration constant $\tilde{d}$ (up to sign) into c and d. Hence ϕ attains the desired form

$$\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2} \tag{5.68}$$

for all $s \in K$. Once again, by Lemma 5.3.11, we can extend the solution by continuity to all relevant values of s. In other words, ϕ has the form (5.68) for all relevant values of s, as desired. In principle, the constants c, d, m and the sign $\pm$ could have different values on different connected components of $s(\mathcal{A})$, but since $\mathcal{A}$ is assumed to be connected, $s(\mathcal{A})$ is also connected and hence the constants and the sign are actually fixed over all of $s(\mathcal{A})$.

Step 4) The condition on $\mathring{\nabla}_i b_j$
Finally we substitute the form (5.65) of ψ and the relation between ρ, σ and λ into the Berwald condition (5.50), leading to

$$y^j \mathring{\nabla}_i b_j = \Big(\lambda|b|^2\beta b_i + 2\rho\beta b_i + \sigma|b|^2 y_i\Big) + \epsilon\lambda\tilde{c}\beta b_i, \tag{5.69}$$

from which we can infer by differentiation that for $\mathring{\nabla}_i b_j$ is symmetric under $i \leftrightarrow j$. This implies that the 1-form β is closed and hence, by Proposition 5.3.9, that the desired Berwald condition holds. This completes the proof. □

Remark 5.3.13. *As a consistency check for the last part of the proof, we may also derive the desired Berwald condition directly from (5.69). We note that the definitions of c and m in terms of $\tilde{c}$ and η above may be inverted as $\tilde{c} = c(1+m)/d$ and $\eta = -mc/d$. Using also the relations between λ, ρ, σ, we find that (5.50) turns*

into

$$y^j \mathring{\nabla}_i b_j = \left(\lambda|b|^2\beta b_i + 2\beta b_i + \sigma|b|^2 y_i\right) + \epsilon\lambda\tilde{c}\beta b_i \tag{5.70}$$

$$= \left(\lambda|b|^2\beta b_i + 2\epsilon\eta\lambda\beta b_i - \epsilon\eta\lambda|b|^2 y_i\right) + \epsilon\lambda\frac{c(1+m)}{d}\beta b_i \tag{5.71}$$

$$= \lambda\left[\left(|b|^2 + 2\epsilon\eta + \epsilon\frac{c(1+m)}{d}\right)\beta b_i - \epsilon\eta|b|^2 y_i\right] \tag{5.72}$$

$$= \lambda\left[\left(|b|^2 - \frac{2\epsilon mc}{d} + \frac{\epsilon c(1+m)}{d}\right)\beta b_i + \frac{\epsilon mc}{d}|b|^2 y_i\right] \tag{5.73}$$

$$= \lambda\left[\left(|b|^2 + \frac{\epsilon c}{d}(1-m)\right)\beta b_i - \epsilon\frac{-mc}{d}|b|^2 y_i\right], \tag{5.74}$$

from which we can infer by differentiation that

$$\mathring{\nabla}_i b_j = \frac{\epsilon\lambda}{d}\left\{\left[\epsilon d|b|^2 + c\,(1-m)\right] b_i b_j + mc|b|^2 a_{ij}\right\}, \tag{5.75}$$

which is precisely (5.48) *provided we identify* $p = \epsilon\lambda/d$.

We now turn to an explicit nontrivial example of a class of Berwald spaces that will turn out to be more general than it may seem, as we will see shortly in Proposition 5.3.15. We will come back to the special case where F is m-Kropina (i.e. $c = 1, d = 0$) in detail in the next chapter. In what follows we will use the convention that indices $i, j \ldots$ run from 1 to n, whereas indices $a, b \ldots$ run from 3 to n.

Example 5.3.14 (A family of generalized m-Kropina spaces of Berwald type)**.** *Consider a generalized m-Kropina metric*

$$F = \alpha\phi(\beta/\alpha), \qquad \phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}, \tag{5.76}$$

with $m \neq 1$, α given by

$$a = -2\mathrm{d}u\mathrm{d}v + H(u,v,x)\,\mathrm{d}u^2 + 2W_a(u,x)\,\mathrm{d}x^a\mathrm{d}u + h_{ab}(u,x)\,\mathrm{d}x^a\mathrm{d}x^b, \tag{5.77}$$

and 1-form $\beta = \mathrm{d}u$, expressed in coordinates $(u, v, x) \coloneqq (u, v, x^3, \ldots, x^n)$, where H, W_a, $a = 3, \ldots n$ are arbitrary smooth functions and h_{ab} is a pseudo-Riemannian metric of dimension $n-2$. In the special case that h_{ab} and hence a_{ij} is Lorentzian, metrics of the form (5.77) *are known as Kundt metrics [77, 78, 79], see also Appendix A.1. The 1-form β is null w.r.t. α and the only (possibly) nonvanishing*

component of the covariant derivative of β is:

$$\mathring{\nabla}_u b_u = -\tfrac{1}{2}\partial_v H\,. \tag{5.78}$$

Hence condition (5.39) is satisfied with $|b|^2 = 0$ and $p = -\partial_v H/2c(1-m)$ and hence this geometry is Berwald. The affine connection is determined by the tensor T^k_{ij}, which is of the form (5.46) with coefficients given by

$$\lambda = \frac{-\epsilon d}{2c(1-m)}\partial_v H, \qquad \rho = -\sigma = \frac{m}{2(1-m)}\partial_v H. \tag{5.79}$$

In fact, as shown by the following proposition, this example exhausts essentially all possible generalized m-Kropina spaces of Berwald type with closed null 1-form.

Proposition 5.3.15. *Let $F = \alpha\phi(\beta/\alpha)$, $\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}$ be a generalized m-Kropina space with $n > 2$, $m \neq 1$ and suppose that β is a closed null 1-form. TFAE:*

(i) F is Berwald

(ii) $\nabla_i b_j = p\, b_i b_j$ for a smooth function p;

(iii) F is given, locally, by the Finsler metric of Example 5.3.14;

In this case $p = -\frac{1}{2}\partial_v H$ and

$$\Gamma^\ell_{ij} = {}^\alpha\Gamma^\ell_{ij} + \frac{mp}{1-m}\left(a_{ij}b^\ell - \delta^\ell_j b_i - \delta^\ell_i b_j\right). \tag{5.80}$$

Proof. Setting $|b|^2 = 0$ in Proposition 5.3.9 it follows that F is Berwald iff there exists a smooth function p on M such that $\mathring{\nabla}_i b_j = pc(1-m)b_i b_j$. Since $c \neq 0$ by Lemma 5.3.7, and since we have assumed that $(1-m) \neq 0$, this is equivalent to condition (ii), proving the equivalence between (i) and (ii). The implication (iii)⇒(ii) is provided by Example 5.3.14, and finally, the implication (ii)⇒(iii) follows by Lemma 5.3.16 below. □

In contrast to most of the results in the chapter, Proposition 5.3.15 as well as Lemma 5.3.16 below are based on [H3] rather than [H1].

Lemma 5.3.16. *Let $n > 2$ and let a be a pseudo-Riemannian metric and b a nowhere vanishing 1-form. If $|b|^2 = 0$ and $\mathring{\nabla}_i b_j = p\, b_i b_j$ for some function p then*

around each point in M there exist local coordinates $(u, v, x^3, \dots, x^n)$ such that $b = \mathrm{d}u$ and

$$a = -2\mathrm{d}u\mathrm{d}v + H(u, v, x)\mathrm{d}u^2 + 2W_a(u, x)\mathrm{d}u\mathrm{d}x^a + h_{ab}(u, x)\mathrm{d}x^a\mathrm{d}x^b, \tag{5.81}$$

with h some pseudo-Riemannian metric of dimension $n-2$, and p is then given by $p = -\frac{1}{2}\partial_v H$.

Proof. Because $\mathring{\nabla}_i b_j$ is symmetric, b is closed and hence locally exact, and, since $|b|^2 = 0$, by assumption, it follows by Proposition A.1.1 that the metric a can be written in the form

$$a = -2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_b\mathrm{d}u\mathrm{d}x^b + h_{bc}\mathrm{d}x^b\mathrm{d}x^c. \tag{5.82}$$

It remains to be shown that the functions W_a and h_{ab} do not depend on coordinate v. To this end, we employ on the one hand the condition $\mathring{\nabla}_i b_j = p\, b_i b_j = p\, \delta_i^u \delta_j^u$ expressed in our preferred coordinates, and on the other hand the expression for $\mathring{\nabla}_i b_j$ computed explicitly in the preferred coordinates. Using the fact that $b_i = \delta_i^u$ and $a^{ui} = -\delta_v^i$ and $a_{iv} = 0$, we find after a straightforward computation that the latter is given by $\mathring{\nabla}_i b_j = -\frac{1}{2}\frac{\partial a_{ij}}{\partial v}$, thus leading to the requirement that

$$p\, \delta_i^u \delta_j^u = -\frac{1}{2}\frac{\partial a_{ij}}{\partial v}. \tag{5.83}$$

Plugging in the explicit form of a_{ij} in our preferred coordinates, this leads to

$$p = -\tfrac{1}{2}\partial_v H, \qquad \partial_v W_a = \partial_v h_{ab} = 0, \tag{5.84}$$

as desired. □

As a final remark, we note that some of the results obtained in this chapter for Berwald spaces can easily be generalized to so-called *generalized* Berwald spaces, defined by the property that they admit a (not necessarily torsion-free) metric-compatible affine connection. The only difference is that in that case, T^k_{ij} need not be symmetric in $i \leftrightarrow j$, as it has to be for Berwald spaces. We note, however, that we have applied *Christoffel's trick* in some of the proofs, which is based precisely on this symmetry, so these parts will not necessarily generalize immediately.

CHAPTER 6

Pseudo-Riemann Metrizability

Given a Finsler space of Berwald type, the Cartan nonlinear connection defines a linear connection on TM that is, by definition, torsion-free and metric-compatible. The natural question thus arises whether there exists a pseudo-Riemannian metric that has this connection as its Levi-Civita connection. Simply put, is (the affine connection on) every Berwald space metrizable? For positive definite Finsler spaces defined on $\mathcal{A} = TM_0$ this is indeed the case, as was first shown by Szabo.

Theorem 6.0.1 (Szabo's Metrization Theorem [42])**.** *Any positive definite Berwald space with $\mathcal{A} = TM_0$ is metrizable by a Riemannian metric.*

The proofs of this theorem [42, 80] rely on procedures such as averaging [81] over the indicatrix $S_x = \{(x, y) \in T_xM \,:\, F(x, y) = 1\}$ for which it is essential that the Finsler metric F is sufficiently smooth and defined everywhere on $\mathcal{A} = TM_0$. In the case of Finsler metrics of indefinite signature, however, the domain where F is defined is typically only a proper subset of TM_0 and hence the proofs do not extend to this case. And even if $\mathcal{A} = TM_0$, the fact that the indicatrix is not compact in indefinite signatures poses problems.

It was shown in [H2] that Szabo's metrization theorem is indeed not valid for Finsler spaces in the general sense of our Definition 1.2.3, and in [H3] the situation was further investigated and necessary and sufficient conditions for (local) metrizability were obtained in the specific case of m-Kropina spaces with closed null 1-form. In this chapter we establish these results and expand on

them considerably, covering not only m-Kropina metrics but also generalized m-Kropina metrics, and arbitrary 1-forms.

The culprit behind all known counterexamples to Szabo's theorem is the fact that the affine Ricci tensor is in general not symmetric. This property, symmetry of the affine Ricci tensor, is clearly a necessary condition for metrizability, and we will see that for all specific metrics investigated here, this is in fact also a sufficient condition, at least locally. In other words, if the affine Ricci tensor is symmetric then any point in M has a neighborhood on which the geometry is metrizable. This property is called *local metrizability*, and it will be the main focus of this chapter. In what follows we will thus use the following definition.

Definition 6.0.2. *A Berwald space with underlying manifold M is said to be locally metrizable if each point in M has a neighborhood that admits a pseudo-Riemannian metric whose Levi-Civita connection coincides with the affine connection restricted to that neighborhood.*

After establishing the necessary and sufficient conditions for local metrizability, we classify all locally metrizable and Ricci-flat m-Kropina metrics in (1+3)D, and we discuss the implications for vacuum solutions to the field equations in Finsler gravity. But before we turn to these specific types of spaces, we give the general argument for nonmetrizability, together with a simple explicit counterexample.

6.1 Nonmetrizability of Berwald spaces

We start with an adaptation of the results developed in [H2]. A simple way to see why not all Berwald spacetimes are metrizable is to consider the lack of symmetry of the affine Ricci tensor $\bar{R}_{ij}$. Recall that the affine Ricci tensor is given by (4.4) in terms of the affine connection induced by the canonical nonlinear connection. Recall further that the curvature tensor of the affine connection of a Berwald space is related to the nonlinear curvature via (4.3). Finally, define the function $f = \ln\sqrt{|\det g|}$ and note that

$$\delta_i f = \frac{1}{2|\det g|}\delta_i|\det g| = \frac{1}{2}g^{kl}\delta_i g_{kl} = \Gamma^k{}_{ik}, \tag{6.1}$$

$$\bar{\partial}_i f = \frac{1}{2|\det g|}\bar{\partial}_i|\det g| = \frac{1}{2}g^{kl}\bar{\partial}_i g_{kl} = g^{kl}C_{ikl} = C_i\,. \tag{6.2}$$

Then, using the definition of the nonlinear curvature (2.14) in terms of commutators of horizontal vector fields and the fact that $\delta_i\Gamma^k_{jl} = \partial_i\Gamma^k_{jl}$, since $\bar{\partial}_i\Gamma^k_{jl} = 0$,

we find that

$$\bar{R}_{ij} - \bar{R}_{ji} = \delta_i \Gamma^k{}_{jk} - \delta_j \Gamma^k{}_{ik} = \delta_i \delta_j f - \delta_j \delta_i f \tag{6.3}$$
$$= [\delta_i, \delta_j] f = R^k{}_{ij} \bar{\partial}_k f = \bar{R}^k{}_{lij} y^l \bar{\partial}_k f \tag{6.4}$$
$$= \bar{R}^k{}_{lij} y^l C_k. \tag{6.5}$$

Hence the skew-symmetric part of the affine Ricci tensor is given by

$$\bar{R}_{[ij]} = \tfrac{1}{2} \bar{R}^k{}_{lij} y^l C_k. \tag{6.6}$$

If the Finsler Lagrangian is pseudo-Riemannian then $C_k = 0$ and $\bar{R}_{ij} = \bar{R}_{ji}$, but for general Finsler metrics of Berwald type the RHS of (6.6) need not vanish. We will give an explicit example below. In such a case it is immediately clear that the affine connection cannot be obtained as the Levi-Civita connection of a pseudo-Riemannian metric, because the Levi-Civita connection always has a symmetric Ricci tensor. This shows that not all Berwald spaces are pseudo-Riemann metrizable, as was demonstrated originally in [H2].

Example 6.1.1 (Counterexample to Szabó's theorem)**.** *Let the Lorentzian metric α and the 1-form β be given by*

$$a = -2\,\mathrm{d}u\mathrm{d}v + v\,\psi(x,y)\,\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2, \qquad \beta = \mathrm{d}u, \tag{6.7}$$

where ψ is any scalar function that is not constant, and let F be any generalized m-Kropina metric $F = \alpha\phi(\beta/\alpha)$, $\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2}$ constructed from these building blocks. Note that these form a subclass of the family of Berwald metrics described in Example 5.3.14. A straightforward calculation (or application of Proposition 6.2.2 below) shows that

$$\tfrac{1}{2}\left(\bar{R}_{ux} - \bar{R}_{xu}\right) = \tfrac{m}{m-1}\partial_x \psi, \qquad \tfrac{1}{2}\left(\bar{R}_{uy} - \bar{R}_{yu}\right) = \tfrac{m}{m-1}\partial_y \psi. \tag{6.8}$$

Since ψ is not constant, it is clear that unless $m = 0$, the affine Ricci tensor is not symmetric and hence the affine connection of F does not coincide with the Levi-Civita connection of any pseudo-Riemannian metric.

The culprit behind the fact that $\bar{R}_{ij}$ fails to be symmetric in general, and indeed the culprit behind all currently known counterexamples to Szabo's theorem, seems to be not the signature, but the fact that $\mathcal{A} \subsetneq TM_0$ is a proper subset. This is shown by the following theorem, the proof of which uses the properties of the (positively) projective tangent space at a point $x \in M$, defined as the quotient $PT_xM^+ \equiv (T_xM \setminus 0)/\!\sim$ where $y \sim y'$ iff $y' = \lambda y$ for some $\lambda > 0$. This

construction is similar to that of standard projective spaces, see e.g. [54, 82]; for details we refer to [38] and references therein. The two essential properties that we will use in the proof below are that (i) PT_xM^+ is compact, and that (ii) any smooth 0-homogeneous function on $T_xM \setminus 0$ can be identified with a smooth function $\hat{f}$ on PT_xM^+ via $\hat{f}([y]) = f(y)$, where $[y]$ is the equivalence class of $y \in T_xM \setminus 0$ under the equivalence relation $\sim$.

Theorem 6.1.2. *For a Berwald space with $\mathcal{A} = TM_0$, the affine Ricci tensor is symmetric, $R_{ij} = R_{ji}$.*

Proof. We employ the expression $R_{ij} - R_{ji} = R^k{}_{ij}\bar{\partial}_k f$ derived above for the skew-symmetric part of the Ricci tensor and we start by noting that the LHS only depends on x, not y, for Berwald spaces. It thus suffices to show that for each $x \in M$ there exists at least one $y \in \mathcal{A} = TM_0$ such that $R^k{}_{ij}\bar{\partial}_k f = 0$; this then implies that the expression vanishes on all of TM_0. So let us fix some $x \in M$. Since $y \mapsto f(x, y)$ is smooth and homogeneous of degree 0 it can be interpreted as a smooth function $\hat{f}$ on PT_xM^+. Since $\hat{f}$ is in particular continuous and PT_xM^+ is compact, $\hat{f}$ attains its minimum on PT_xM^+. Suppose this minimum is attained at the equivalence class $[y] \in PT_xM^+$. That means that the original function $y \mapsto f(x, y)$ attains its minimum at $y \in T_xM \setminus 0$, so that $\bar{\partial}_i f|_{(x,y)} = 0$ and hence $R^k{}_{ij}\bar{\partial}_k f|_{(x,y)} = 0$. Since x was arbitrary this completes the proof. □

6.2 Metrizability of generalized m-Kropina spaces with closed null 1-form

This section focuses on generalized m-Kropina metrics with closed null 1-form (i.e. $\mathrm{d}b = 0$ and $|b|^2 = a_{ij}b^i b^j = 0$) of Berwald type. We will also assume throughout this section that $n := \dim M > 2$ and that $c \neq 0$ and $m \neq 1$ and we remind the reader that, since $\mathcal{A}$ is assumed to be contained in the subset of TM characterized by $s > 0$, $c + ds^2 > 0$, in order for any real power of those expressions to be defined, we may assume without loss of generality that b_i is nowhere vanishing. Furthermore, we will stick to our convention that indices $a, b, c, \ldots$ run from 3 to n, whereas indices $i, j, k, \ldots$ run from 1 to n. Many results in this section have been published in [H3], but only with regard to *standard* m-Kropina spaces. It turns out that exactly the same analysis holds up for *generalized* m-Kropina spaces, and we present this extended analysis here for the first time.

Assuming our generalized m-Kropina space is Berwald it follows from Proposition 5.3.15 that in suitable coordinates we have $\beta = du$ and

$$a = -2\mathrm{d}u\mathrm{d}v + H(u,v,x)\,\mathrm{d}u^2 + 2W_a(u,x)\,\mathrm{d}x^a\mathrm{d}u + h_{ab}(u,x)\,\mathrm{d}x^a\mathrm{d}x^b, \tag{6.9}$$

and that

$$\Gamma^\ell_{ij} = {}^\alpha\Gamma^\ell_{ij} + \frac{mp}{1-m}\left(a_{ij}b^\ell - \delta^\ell_j b_i - \delta^\ell_i b_j\right) \tag{6.10}$$

with $p = -\frac{1}{2}\partial_v H$. Substituting the form of p into (6.10) and using that $b_i = \delta^u_i$ and hence that $b^\ell = a^{\ell k}b_k = a^{\ell k}\delta^u_k = a^{\ell u} = -\delta^\ell_v$, we obtain the following.

Lemma 6.2.1. *In the coordinates $(u, v, x^3, \dots, x^n)$, the affine connection coefficients can be expressed in terms of the Levi-Civita Christoffel symbols $\mathring{\Gamma}^k_{ij}$ of the pseudo-Riemannian metric α as*

$$\Gamma^k_{ij} = \mathring{\Gamma}^k_{ij} + \Delta\Gamma^k_{ij}, \qquad \Delta\Gamma^k_{ij} = \frac{m}{2(1-m)}\partial_v H\left(a_{ij}\delta^k_v + \delta^k_j\delta^u_i + \delta^k_i\delta^u_j\right). \tag{6.11}$$

We can use the preceding results to analyze the (deviation from the) symmetry of the affine Ricci tensor, which has a very simple expression in these coordinates, as the following result shows.

Proposition 6.2.2. *In the coordinates $(u, v, x^3, \dots, x^n)$, the skew-symmetric part of the affine Ricci tensor is given by*

$$\bar{R}_{[ij]} = -\frac{mn}{4(1-m)}(\delta^u_i\partial_j\partial_v H - \delta^u_j\partial_i\partial_v H). \tag{6.12}$$

Proof. From the definition (4.4) of the affine Ricci tensor of a Berwald space and the fact that $\Gamma^k_{ij} = \Gamma^k_{ji}$ it follows that the skew-symmetric part of the affine Ricci tensor can be written as

$$\bar{R}_{[ij]} \equiv \tfrac{1}{2}\left(\bar{R}_{ij} - \bar{R}_{ji}\right) = \partial_{[i}\Gamma^k_{j]k}. \tag{6.13}$$

Since the Levi-Civita connection $\mathring{\Gamma}^k_{ij}$ has a symmetric Ricci tensor, it follows that $\partial_{[i}\mathring{\Gamma}^k_{j]k} = 0$ and hence we have

$$\bar{R}_{[ij]} = \partial_{[i}\Gamma^k_{j]k} = \partial_{[i}\Delta\Gamma^k_{j]k} = \tfrac{1}{2}\left(\partial_i\Delta\Gamma^k_{kj} - \partial_j\Delta\Gamma^k_{ki}\right). \tag{6.14}$$

Using Lemma 6.2.1 we find that

$$\Delta\Gamma^k_{kj} = \frac{m}{2(1-m)}\partial_v H\left(a_{vj} + \delta^u_j + n\delta^u_j\right) = \frac{mn}{2(1-m)}\partial_v H\delta^u_j, \tag{6.15}$$

where we have used that $a_{vj} = -\delta^u_j$, and hence we obtain

$$\bar{R}_{[ij]} = \frac{mn}{4(1-m)}\left(\delta^u_j\partial_i\partial_v H - \delta^u_i\partial_j\partial_v H\right). \tag{6.16}$$

□

Now we will prove the main result of this section, characterizing the local metrizability of generalized m-Kropina spaces of Berwald type with closed null 1-form.

Theorem 6.2.3. *Let F be a generalized m-Kropina metric of Berwald type with closed null 1-form and with $m \neq 0$ and $\dim M > 2$. The following are equivalent:*

(i) F is locally metrizable;

(ii) The affine Ricci tensor is symmetric, $\bar{R}_{ij} = \bar{R}_{ji}$;

(iii) Around each point in M there exist coordinates $(u, v, x^3, \ldots, x^n)$ such that $b = \mathrm{d}u$ and

$$a = -2\mathrm{d}u\mathrm{d}v + \left[\tilde{H}(u,x) + \rho(u)v\right]\mathrm{d}u^2 + 2W_a(u,x)\mathrm{d}u\mathrm{d}x^a + h_{ab}(u,x)\mathrm{d}x^a\mathrm{d}x^b, \tag{6.17}$$

with $\tilde{H}, \rho, W_3, \ldots W_n$ smooth functions and h a pseudo-Riemannian metric of dimension $n-2$.

In this case, the affine connection restricted to the chart corresponding to the coordinates $(u, v, x^3, \ldots x^n)$ is metrizable by the following pseudo-Riemannian metric:

$$\tilde{a} = e^{\frac{m}{1-m}\int^u \rho(\tilde{u})\mathrm{d}\tilde{u}}a \tag{6.18}$$

Before presenting the proof, we note that if $\rho = 0$ then $\tilde{a} = a$, i.e. the affine connection is metrizable by the defining pseudo-Riemannian metric α. This was to be expected, since in that case the 1-form β is parallel with respect to α, and hence the result follows by Corollary 5.2.5.

Proof. (i) trivially implies (ii). For (ii)⇒(iii) we use the preferred coordinates $(u, v, x^3, \ldots, x^n)$. By Proposition 6.2.2, the only nonvanishing components of the

skew-symmetric part of the affine Ricci tensor are

$$\bar{R}_{[uj]} = -\frac{mn}{4(1-m)}\partial_j\partial_v H, \qquad j = 2, \dots, n. \tag{6.19}$$

By assumption, the Ricci tensor is symmetric, so all of these must vanish. The uv component yields $\partial_v^2 H = 0$ and the remaining components yield $\partial_v\partial_a H = 0$, $a = 3, \dots n$. In other words, H must be linear in v and the corresponding linear coefficient can depend only on the coordinate u. That is,

$$a = -2\mathrm{d}u\mathrm{d}v + \left[\tilde{H}(u,x) + \rho(u)v\right]\mathrm{d}u^2 + 2W_a(u,x)\mathrm{d}u\mathrm{d}x^a + h_{ab}(u,x)\mathrm{d}x^a\mathrm{d}x^b. \tag{6.20}$$

This proves $(ii) \Rightarrow (iii)$. For the final implication $(iii) \Rightarrow (i)$, recall from Lemma 6.2.1 that the affine connection coefficients can be expressed as

$$\Gamma^k_{ij} = \mathring{\Gamma}^k_{ij} + \frac{m}{2(1-m)}\partial_v H\left(a_{ij}\delta^k_v + \delta^k_j\delta^u_i + \delta^k_i\delta^u_j\right). \tag{6.21}$$

On the other hand, an elementary calculation shows that the Levi-Civita Christoffel symbols of the pseudo-Riemannian metric $\tilde{a} = e^{\psi(u)}a$ can be expressed in terms of the original Levi-Civita Christoffel symbols corresponding to a as

$${}^{\tilde{a}}\Gamma^k_{ij} = \mathring{\Gamma}^k_{ij} + \tfrac{1}{2}\psi'(u)\left(a_{ij}\delta^k_v + \delta^k_j\delta^u_i + \delta^k_i\delta^u_j\right). \tag{6.22}$$

In this case, $\psi(u) = \frac{m}{1-m}\int^u \rho(\tilde{u})\mathrm{d}\tilde{u}$, so $\psi'(u) = \frac{m}{1-m}\rho(u) = \frac{m}{1-m}\partial_v H$, and hence it is clear that (6.21) and (6.22) coincide. This shows that the connection coefficients of $\tilde{a}$ coincide with the affine connection coefficients of F, and hence F is locally metrizable, i.e. (i) holds. This completes the proof of the theorem. □

Theorem 6.2.3 provides necessary and sufficient conditions for a generalized m-Kropina space with closed null 1-form to be locally metrizable. Next, we apply the theorem to an explicit example from the physics literature.

Example 6.2.4. *Consider the Finsler VSI spacetimes presented in [44], with the 4-dimensional Finsler metric*

$$F = \left|-2\mathrm{d}u\mathrm{d}v + (\tilde{H} + \rho\, v)\mathrm{d}u^2 + 2W_a\mathrm{d}u\mathrm{d}x^a + \delta_{ab}\mathrm{d}x^a\mathrm{d}x^b\right|^{\frac{1+m}{2}}(\mathrm{d}u)^{-m}, \tag{6.23}$$

where the metric functions $\tilde{H}, \rho, W_a$ depend only on u and x^a, not on v. The way to interpret the RHS is to view all 1-forms as real-valued functions on TM,

products, sums, and powers of which are defined pointwise. The name 'Finsler VSI spacetimes' stems from the fact that the pseudo-Riemannian part α is a VSI spacetime, i.e. it has vanishing scalar curvature invariants [83]. F is an m-Kropina metric and by Proposition 5.3.15, it is of Berwald type. By Theorem 6.2.3 it is locally metrizable if and only if $\partial_a \rho = 0$. The case $\rho = 0$ provides a Finsler version of the gyratonic pp-wave metric [84, 85], which is trivially metrizable by the Lorentzian gyratonic pp-wave metric itself.

As a special case of Example 6.2.4 we obtain a particularly simple, nontrivial example of a locally metrizable Finsler metric.

Example 6.2.5. *Let F be given by (6.23) with $\tilde{H}(u,x) = 0$, $\rho(u,x) = u$ and $W_a(u,x) = 0$. Relabeling the two coordinates x^a as x and y, this leads to the Finsler metric*

$$F = \left|-2\mathrm{d}u\mathrm{d}v + u\,v\,\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2\right|^{\frac{1+m}{2}} (\mathrm{d}u)^{-m}\,, \tag{6.24}$$

which has an affine connection given by the following nonvanishing Christoffel symbols:

$$\Gamma^u_{uu} = \frac{1+m}{1-m}\frac{u}{2}, \qquad \Gamma^v_{uu} = \left(\frac{u^2}{1-m} - 1\right)\frac{v}{2}, \tag{6.25}$$

$$\Gamma^v_{uv} = -\frac{u}{2}, \qquad \Gamma^v_{xx} = \Gamma^v_{yy} = \Gamma^x_{ux} = \Gamma^y_{uy} = \frac{m}{1-m}\frac{u}{2}. \tag{6.26}$$

As indicated by (6.18) in Theorem 6.2.3 this connection is metrizable by the Lorentzian metric

$$\tilde{g} = e^{\frac{mu^2}{2(1-m)}}\left(-2\mathrm{d}u\mathrm{d}v + u\,v\,\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2\right). \tag{6.27}$$

6.3 Metrizability of m-Kropina spaces

The results of the previous section have been obtained using the specific coordinate system introduced in Example 5.3.14 and Proposition 5.3.15, which presupposes that the 1-form β is closed and null. In this section we consider general 1-forms β and so the methods used in the previous section cannot be generalized straightforwardly. However, using different methods it is possible to obtain some interesting results on the metrizability of (standard) m-Kropina spaces with arbitrary 1-form, without the need for introducing a special coordinate system. As a special case, this will reproduce the results obtained above

for m-Kropina spaces with a closed null 1-form, providing an alternative proof of those results. The results in this section were obtained only recently and have not been published yet; we present them here for the first time. In contrast to the results obtained in Section 6.2 they do not generalize in an obvious way to *generalized* m-Kropina spaces.

Hence throughout this section, we will assume F is an m-Kropina metric. Again we restrict our attention to the case $n = \dim M > 2$. Recall from Corollary 5.3.10 that an m-Kropina metric is of Berwald type if and only if there exists a smooth vector field f^i on M such that

$$\mathring{\nabla}_j b_i = m(f_k b^k)a_{ij} + b_i f_j - m f_i b_j, \tag{6.28}$$

and if so, the affine connection is given by

$$\Gamma^\ell_{ij} = \mathring{\Gamma}^\ell_{ij} + \Delta\Gamma^k_{ij}, \qquad \Delta\Gamma^k_{ij} = m a^{\ell k}\left(a_{ij}f_k - a_{jk}f_i - a_{ki}f_j\right) \tag{6.29}$$

in terms of the Levi-Civita connection $\mathring{\Gamma}^\ell_{ij}$ of α. We start, in analogy with the previous section, by deriving a precise formula for the skew-symmetric part of the affine Ricci tensor, which we will denote by $\mathcal{A}(\bar{R})$, i.e. $\mathcal{A}(\bar{R})_{ij} = \bar{R}_{[ij]}$. Moreover, at the risk of stating the obvious, we remark that the expression $\mathrm{d}f$ below denotes the exterior derivative of the 1-form f with components $f_i = a_{ij}f^j$ appearing in the Berwald condition (6.28).

Lemma 6.3.1. *The skew-symmetric part of the affine Ricci tensor is given by*

$$\mathcal{A}(\bar{R}) = -\frac{mn}{2}\mathrm{d}f \tag{6.30}$$

Proof. In general we have $\bar{R}_{[ij]} = \frac{1}{2}(\partial_i \Delta\Gamma^k_{kj} - \partial_j \Delta\Gamma^k_{ki})$ as shown in the proof of Proposition 6.2.2 (see (6.14) in particular). Employing the expression (6.29) for $\Delta\Gamma^k_{ij}$, we find that

$$\Delta\Gamma^k_{kj} = -mnf_j \tag{6.31}$$

and this yields

$$\bar{R}_{[ij]} = \frac{mn}{2}(\partial_j f_i - \partial_i f_j), \tag{6.32}$$

which is precisely the coordinate expression of (6.30), as desired. □

We then have the following characterization of local metrizability.

Theorem 6.3.2. *For an m-Kropina metric $F = \alpha^{1+m}\beta^{-m}$ of Berwald type with* $\dim M > 2$*, the following are equivalent:*

(i) F is locally metrizable;

(ii) The affine Ricci tensor is symmetric, $\bar{R}_{ij} = \bar{R}_{ji}$;

(iii) f_i is a closed 1-form.

In that case we can locally write $f_i = \partial_i\psi$ for some $\psi \in C^\infty(M)$ and we may define $\tilde{a}_{ij} = e^{-2m\psi}a_{ij}$ and $\tilde{b}_i = e^{-(1+m)\psi}b_i$ and define $\tilde{\alpha}$ and $\tilde{\beta}$ accordingly. Then F can be written as $F = \tilde{\alpha}^{1+m}\tilde{\beta}^{-m}$, $\tilde{\beta}$ is covariantly constant with respect to $\tilde{\alpha}$, and $\tilde{a}_{ij}$ provides a local metrization of F.

Proof. The first two implications are rather straightforward. (i)⇒(ii): If F is locally metrizable then it follows trivially that the Ricci tensor is symmetric. (ii)⇒(iii): If the Ricci tensor is symmetric then Lemma 6.3.1 shows that $\mathrm{d}f = 0$. The remaining nontrivial part of the proof is showing that (iii) implies (i). (iii)⇒(i): If f_i is closed then f is locally exact, i.e. $f_i = \partial_i\psi$ for some (locally defined) function ψ on M. Consider the following transformed metric and 1-form, $\tilde{a}_{ij} = e^{-2m\psi}a_{ij}$ and $\tilde{b}_i = e^{-(1+m)\psi}$ and define $\tilde{\alpha}$ and $\tilde{\beta}$ accordingly. Then $F = \alpha^{1+m}\beta^{-m} = \tilde{\alpha}^{1+m}\tilde{\beta}^{-m}$ and it is straightforward to show that the Christoffel symbols of $\tilde{a}_{ij}$ are given by

$$\tilde{\Gamma}^k_{ij} = \mathring{\Gamma}^k_{ij} - m\left(f_i\delta^k_j + f_j\delta^k_i - f^k a_{ij}\right) \tag{6.33}$$

in terms of those of a_{ij}. Hence if we denote by $\tilde{\nabla}$ the Levi-Civita connection of $\tilde{\alpha}$ then we find that

$$\tilde{\nabla}_j\tilde{b}_i = e^{-(1+m)\psi}\left[\mathring{\nabla}_j b_i - (m(f_k b^k)a_{ij} + b_i f_j - m f_i b_j)\right] = 0, \tag{6.34}$$

which vanishes, by (6.28). In other words, $\tilde{b}_i$ is covariantly constant w.r.t. $\tilde{a}_{ij}$, and since we can write F as $F = \tilde{\alpha}^{1+m}\tilde{\beta}^{-m}$, it follows from Corollary 5.2.2 that the affine connection on F is just the Levi-Civita connection of $\tilde{a}_{ij}$. Hence F is locally metrizable, namely by $\tilde{a}_{ij}$, completing the proof. □

The mapping $(\alpha, \beta) \mapsto (\tilde{\alpha}, \tilde{\beta})$ is a generalization of what is called the *f-change* in [73, Chapter 6].

We can now distinguish several cases.

1. b is not null. In this case, f_i is necessarily closed, which follows immediately from Proposition 5.3.8 and which was also observed in [73]. Hence we have the following.

Corollary 6.3.3. *An m-Kropina space with $|b|^2 \neq 0$ is locally metrizable.*

2. If b is null and closed then Section 6.2 completely characterizes local metrizability in terms of local coordinates.

3. If b is null but not necessarily closed then the local metrizability of F can also be characterized in local coordinates, as will be discussed next.

Proposition 6.3.4. *If $|b|^2 = 0$ then F is locally metrizable if and only if around each point in M there exist coordinates $(u, v, x^3, \ldots, x^n)$ such that*

$$a = e^{2m\psi}\left(-2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_a\mathrm{d}u\mathrm{d}x^a + h_{ab}\mathrm{d}x^a\mathrm{d}x^b\right), \tag{6.35}$$

$$b = e^{(1+m)\psi}\mathrm{d}u, \tag{6.36}$$

where ψ is any smooth function on M and the metric functions H, W_a, h_{ab} depend only on u and x^a. In that case, F can be written as

$$F = \left|-2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_a\mathrm{d}u\mathrm{d}x^a + h_{ab}\mathrm{d}x^a\mathrm{d}x^b\right|^{(1+m)/2}(\mathrm{d}u)^{-m} \tag{6.37}$$

and the metric $-2\mathrm{d}u\mathrm{d}v+H\mathrm{d}u^2+2W_a\mathrm{d}u\mathrm{d}x^a+h_{ab}\mathrm{d}x^a\mathrm{d}x^b$ provides a local metrization of F.

Proof. We first prove the 'only if' part. By Theorem 6.3.2, if F is locally metrizable then there exist $\tilde{a}_{ij} = e^{-2m\psi}a_{ij}$ and $\tilde{b}_i = e^{-(1+m)\psi}b_i$ with corresponding $\tilde{\alpha}$ and $\tilde{\beta}$ such that $F = \tilde{\alpha}^{1+m}\tilde{\beta}^{-m}$ and such that $\tilde{\beta}$ is covariantly constant with respect to $\tilde{a}_{ij}$. By Proposition A.2.1 (recall that we may WLOG assume that b_i does not vanish by our definition of $\mathcal{A}$) it thus follows that locally, we may write

$$\tilde{a} = -2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_a\mathrm{d}u\mathrm{d}x^a + h_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{6.38}$$

$$\tilde{b} = \mathrm{d}u, \tag{6.39}$$

where the functions H, W_a, h_{ab} depend only on u and x^a. It is immediately verified that a, b and F now attain the desired form. Conversely, if a and b have the stated form, then F is given by (6.37) in terms of $\tilde{a}$ and $\tilde{b}$, and since $\tilde{b} = \mathrm{d}u$ is covariantly constant with respect to $\tilde{\alpha}$, the latter provides a local metrization of F, by Corollary 5.2.2. □

As a special case of Proposition 6.3.4 we recover the main result of the previous section, Theorem 6.2.3, as shown by the following corollary.

Corollary 6.3.5. *If b is null and closed then F is locally metrizable if and only if around each point in M there exist coordinates $(u, v, x^3, \dots, x^n)$ such that*

$$F = \left|-2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_a\mathrm{d}u\mathrm{d}x^a + h_{ab}\mathrm{d}x^a\mathrm{d}x^b\right|^{(1+m)/2}(\mathrm{d}u)^{-m}\,, \tag{6.40}$$

where

$$H = \tilde{H}(u, x^a) + \rho(u)v, \qquad W_a = W_a(u, x^a), \qquad h_{ab} = h_{ab}(u, x^a). \tag{6.41}$$

Proof. Suppose that F is locally metrizable, then a and b are given by (6.35) and (6.36), where the metric functions do not depend on the coordinate v. Furthermore, the fact that the 1-form $b = e^{(1+m)\psi}\mathrm{d}u$ is closed implies that $\psi = \psi(u)$. In this case, it can be verified that the coordinate transformation given by $\bar{u} = \int^u e^{(1+m)\psi(u)}\mathrm{d}u$ and $\bar{v} = e^{(m-1)\psi(u)}v$ brings the 1-form into the form $b = \mathrm{d}\bar{u}$ and the metric into the form

$$a = -2\mathrm{d}\bar{u}\mathrm{d}\bar{v} + \bar{H}\mathrm{d}\bar{u}^2 + 2\bar{W}_a\mathrm{d}\bar{u}\mathrm{d}x^a + \bar{h}_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{6.42}$$

where

$$\bar{H} = e^{-2\psi}\left[H + 2(m-1)\psi' v\right], \quad \bar{W}_a = e^{(m-1)\psi}W_a, \quad \bar{h}_{ab} = e^{2m\psi}h_{ab}. \tag{6.43}$$

Identifying $\tilde{H}(u, x^a) = e^{-2\psi}H$ and $\rho(u) = 2e^{-2\psi}(m-1)\psi'$ it can be verified that F attains the desired form in the new coordinates.

Conversely, if F has the stated form, then we define $\psi(u)$ such that $\rho(u) = 2e^{-2\psi}(m-1)\psi'$, by setting $\psi(u) \equiv -\frac{1}{2}\ln\left|\frac{1}{1-m}\int\rho(u)\mathrm{d}u\right|$, and then we can simply perform the same coordinate transformation backward. It is clear that this results in a metric and 1 form of the form (6.35) and (6.36) and hence, by Proposition 6.3.4, F is Berwald. □

6.4 Ricci-flatness of m-Kropina spaces

We continue our investigation of m-Kropina metrics and turn to Ricci-flatness. Recall that a Finsler space is said to be Ricci-flat if the Finsler-Ricci tensor vanishes, $R_{ij} = 0$, or equivalently, if the Ricci scalar vanishes, $\mathrm{Ric} = 0$. We restrict ourselves in this section to the case where M is four-dimensional, which

is the case that is relevant to Finsler gravity. The results in this section have not been published yet; we present them here for the first time.

6.4.1 Ricci-flat nonnull m-Kropina metrics in (1+3)D

First, we investigate Ricci-flat m-Kropina metrics with $|b|^2 \neq 0$ in $(1+3)$D, by which we mean that the underlying manifold M has dimension 4 and that a_{ij} has Lorentzian signature.

Proposition 6.4.1. *A $(1+3)$D m-Kropina space with $|b|^2 \neq 0$ is Ricci-flat if and only if it is locally of the form*

$$F = \left|\eta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right|^{(1+m)/2}\left(c_i\mathrm{d}x^i\right)^{-m} \tag{6.44}$$

with $\eta_{ij} = diag(-1,1,1,1)$ and $c_i = const.$

Proof. We first prove the 'if' direction. Since $c_i\mathrm{d}x^i$ is covariantly constant with respect to η_{ij}, it follows by Corollary 5.2.2 that any F defined by (6.44) is Ricci-flat, since the Ricci tensor of η_{ij} (which vanishes) must coincide with the affine Ricci tensor of F, which therefore vanishes, implying that also the Finsler-Ricci tensor vanishes, by Lemma 4.1.1.

Conversely, to prove the 'only if' direction, suppose that F is Ricci-flat. Then, defining $\tilde{\alpha}, \tilde{\beta}$ as in Theorem 6.3.2, we may write $F = \tilde{\alpha}^{1+m}\tilde{\beta}^{-m}$ and F has the same affine connection as $\tilde{\alpha}$, hence if F is Ricci-flat then $\tilde{\alpha}$ is Ricci-flat. But $\tilde{\alpha}$ admits a nonnull covariantly constant 1-form, namely $\tilde{b}$, so if $\tilde{\alpha}$ is Ricci-flat it must actually be flat, by Corollary A.2.3, and hence, in suitable coordinates it can be written as $\tilde{a} = \eta_{ij}\mathrm{d}x^i\mathrm{d}x^j$. Since $\tilde{\beta}$ is covariantly constant with respect to this metric, it must have constant coefficients in these coordinates, and the conclusion of the theorem follows. □

We remark that the assumption about the dimensionality is not void, because the conclusion of Corollary A.2.3 is not necessarily true if $\dim M \neq 4$. We also point out that the *representation* of F in terms of a specific choice of α and β is, of course, not unique. However, since any α is related to the $\tilde{\alpha}$ from the proposition by a conformal transformation, we have the following two corollaries.

Corollary 6.4.2. *If F is Ricci-flat and $|b|^2 \neq 0$ then α is conformally flat.*

Corollary 6.4.3. *If α is Ricci-flat and $|b|^2 \neq 0$ then F is Ricci-flat only if α is flat and β has constant components in flat coordinates.*

Proof. If F is Ricci-flat and $|b|^2 \neq 0$ then, by the previous corollary, α is conformally flat as well as Ricci-flat. It is thus a simple consequence of the Ricci-decomposition of the Riemann tensor that α must be flat. From Proposition 6.4.1 it then follows that β must have constant coefficients in flat coordinates. □

These results have important consequences for the construction of solutions to the field equations in Finsler gravity, the topic of Part III. Most (properly Finslerian) solutions currently known in the literature are obtained through constructions where a vacuum solution α in general relativity is deformed into a Finslerian solution F that has the form of an (α, β)-metric. Then the Ricci-flatness of F (which is a sufficient condition to be a vacuum solution in Finsler gravity) is usually derived somehow from the Ricci-flatness of α (the exact condition to be a vacuum solution to Einstein's field equation). For m-Kropina metrics of Berwald type, our results above show that if $|b|^2 \neq 0$ this procedure can only lead to a single possible Finsler gravity vacuum solution[1], namely the metric (6.44). It follows that, apart from this trivial solution, m-Kropina spacetimes only lend themselves to constructing nontrivial Finslerian vacuum solutions in the above way for *null* 1-forms β. We consider those next.

6.4.2 Ricci-flat null m-Kropina metrics in (1+3)D

While an m-Kropina space with $|b|^2 \neq 0$ is always locally metrizable, as we have seen in Corollary 6.3.3, this is not always the case when the 1-form is null. With regard to Ricci-flatness, it turns out that in the $|b|^2 = 0$ case we have to add the additional assumption that F is locally metrizable in order to obtain nice results. Alternatively, one may choose to replace everywhere the two conditions *locally metrizable and Ricci-flat* by the single condition *affinely Ricci-flat*, the latter meaning that the *affine* Ricci tensor vanishes (rather than the Finsler-Ricci tensor, which is its symmetrization, by Lemma 4.1.1). Affine Ricci-flatness is not a very common notion in the literature, though, so we will stick to locally metrizable and Ricci-flat.

Proposition 6.4.4. *Let F be an m-Kropina space. Then*

$$F \text{ is affinely Ricci-flat} \quad \Leftrightarrow \quad F \text{ is locally metrizable and Ricci-flat}$$

Proof. If F is affinely Ricci-flat then it is trivially Ricci-flat, since the Finsler-Ricci tensor is the symmetrization of the affine Ricci tensor. It is also locally

[1] Strictly speaking, this is only true locally. Our results do not say anything about the global topology of the manifold.

metrizable, by Theorem 6.3.2, since the (vanishing) affine Ricci tensor is, in particular, symmetric. For the converse implication, if F is locally metrizable then, by Theorem 6.3.2, its Ricci tensor is symmetric and hence the affine Ricci tensor coincides with the Finsler-Ricci tensor. The vanishing of the Finsler-Ricci tensor thus implies, in this case, the vanishing of the affine Ricci tensor. □

The following lemma will allow us to obtain the $|b|^2 = 0$ analog of Proposition 6.4.1, and it will prove to be useful more generally in Part III.

Lemma 6.4.5. *Let a Lorentzian metric a and 1-form b on a 4-dimensional manifold be given by*

$$a = -2\mathrm{d}u\mathrm{d}v + H(u,x)\,\mathrm{d}u^2 + 2W_a(u,x)\,\mathrm{d}x^a\mathrm{d}u + h_{ab}(u,x)\mathrm{d}x^a\mathrm{d}x^b, \tag{6.45}$$

$$b = \mathrm{d}u. \tag{6.46}$$

If a is Ricci-flat then a and b can be expressed, in suitable coordinates, as

$$a = -2\mathrm{d}\bar{u}\mathrm{d}\bar{v} + \bar{H}(\bar{u},\bar{x},\bar{y})\mathrm{d}\bar{u}^2 + \mathrm{d}\bar{x}^2 + \mathrm{d}\bar{y}^2, \tag{6.47}$$

$$b = \mathrm{d}\bar{u}, \tag{6.48}$$

such that $\bar{H}$ satifies $\partial_{\bar{x}}^2\bar{H} + \partial_{\bar{y}}^2\bar{H} = 0$.

Proof. We first consider only the metric a. Since b is null and covariantly constant, it follows from Proposition A.3.1 that a can be expressed in suitable coordinates as

$$a = -2\mathrm{d}u\mathrm{d}v + H(u,x,y)\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2 \tag{6.49}$$

such that $\delta^{ab}\partial_a\partial_b H = 0$. We are, however, not only interested in the transformation behavior of a alone but also in that of b. In [86], on which Proposition A.3.1 is based, no explicit coordinate transformation is given, so to see why we may assume without loss of generality that $b = \mathrm{d}u$ remains invariant, we will use the properties of the general set of coordinate transformations

$$(u, v, x^1, x^2) \mapsto (\bar{u}, \bar{v}, \bar{x}^1, \bar{x}^2) \tag{6.50}$$

that leave the form of the metric (6.45) invariant (but that generally change the expressions for the metric functions $H, W_a, h_{ab} \mapsto \bar{H}, \bar{W}_a, \bar{h}_{ab}$). Such transformations have been classified [79, Section 31.2] and they all have the specific property that $u = \phi(\bar{u})$ for some function ϕ depending on $\bar{u}$ alone. Hence this applies in particular to the desired transformation that relates (6.45) to (6.49).

Let's be a little bit more specific with the notation. What we have argued so far is that there must exist a coordinate transformation (6.50) such that (6.45) turns into

$$a = -2\mathrm{d}\bar{u}\mathrm{d}\bar{v} + \bar{H}(\bar{u},\bar{x},\bar{y})\mathrm{d}\bar{u}^2 + \mathrm{d}\bar{x}^2 + \mathrm{d}\bar{y}^2 \tag{6.51}$$

and such that $u = \phi(\bar{u})$. The latter implies we may express the 1-form as $b = \mathrm{d}u = \phi'(\bar{u})\mathrm{d}\bar{u}$, or equivalently $\bar{b}_\mu = \phi'(\bar{u})\delta^u_\mu$. The trick is now to realize that since b is covariantly constant (cf. Proposition A.2.1) with respect to a, we must have $\bar{\nabla}_\mu \bar{b}_\nu = 0$ also in the new coordinates. All Christoffel symbols $\bar{\Gamma}^u_{\mu\nu}$ of the metric (6.51) with upper index u vanish identically, however (see (A.13)), and hence ϕ must satisfy

$$0 \overset{!}{=} \bar{\nabla}_{\bar{u}}\bar{b}_{\bar{u}} = \partial_{\bar{u}} b_{\bar{u}} - \bar{\Gamma}^u_{uu}\phi'(\bar{u}) = \phi''(\bar{u}). \tag{6.52}$$

It follows that $\phi'(\bar{u}) = C =$ constant, i.e. $b = C\mathrm{d}\bar{u}$. In this case, it is easily seen that scaling $\bar{u}$ by C and scaling $\bar{v}$ by $1/C$ leaves the form of the metric (6.51) invariant (while changing $\bar{H}$) and brings the 1-form back into its original form $b = \mathrm{d}\bar{u}$. This shows that we may assume without loss of generality that $C = 1$ and hence that the 1-form remains invariant under the coordinate transformation relating (6.45) to (6.49). This completes the proof. □

Proposition 6.4.6. *Any locally metrizable* $(1+3)D$ *m-Kropina space with* $|b|^2 = 0$ *is Ricci-flat if and only it is locally of the form*

$$F = \left|-2\mathrm{d}u\mathrm{d}v + H(u,x)\mathrm{d}u^2 + \delta_{ab}\mathrm{d}x^a\mathrm{d}x^b\right|^{(1+m)/2}(\mathrm{d}u)^{-m} \tag{6.53}$$

such that H *satifies* $\delta^{ab}\partial_a\partial_b H = 0$.

Proof. First of all, since $\mathrm{d}u$ is covariantly constant, it follows by Corollary 5.2.2 that any F defined by (6.53) is Ricci-flat, since the Ricci tensor of the Lorentzian metric (which vanishes, by Proposition A.3.1) must coincide with the affine Ricci tensor of F, which therefore vanishes, implying that also the Finsler-Ricci tensor vanishes, by Lemma 4.1.1.

Conversely, suppose that F is Ricci-flat. Since F is locally metrizable it can be written in the standard form (6.37) with a and b given by (6.45) and (6.46), respectively. Since the 1-form $b = \mathrm{d}u$ is covariantly constant with respect to the corresponding metric a, it follows that the affine connection on F coincides with the affine connection on a. So since F is Ricci-flat, a is Ricci-flat as well. Hence

it follows from Lemma 6.4.5 that F can be written locally in the form (6.53), as desired. □

With this, we have classified (locally, at least) all locally metrizable, Ricci-flat m-Kropina metrics in $(1+3)$D. We end this chapter with a final remark about the implications of our findings for the study of vacuum solutions in Finsler gravity. The results show that any Berwald m-Kropina solution to the Finsler field equations in vacuum that is not of one of the trivial forms (6.44) or (6.53), must be one for which at least one of the following two conditions is satisfied:

(i) the 1-form is null and F is not locally metrizable;

(ii) F is not Ricci-flat.

In light of the fact (cf. Part III) that for Berwald m-Kropina spaces with null 1-form β the vacuum field equations are equivalent to the vanishing of the Finsler-Ricci tensor [44], it follows that we can equivalently write these two conditions as:

(i) the 1-form is null and F is not locally metrizable;

(ii) the 1-form is not null and F is not Ricci-flat.

To complete the classification of exact vacuum solutions to the field equations within the class of Berwald m-Kropina spaces, a better understanding of these two scenarios is required. While it is currently not known whether vacuum solutions satisfying (ii) exist at all, some examples of Ricci-flat spaces satisfying (i) have been obtained, as shown by the example below. Nevertheless, a thorough investigation of both cases (i) and (ii) has yet to be carried out.

Example 6.4.7. *Consider the Finsler metric from Example 6.2.4,*

$$F = \left|-2\mathrm{d}u\mathrm{d}v + (\tilde{H} + \rho\, v)\mathrm{d}u^2 + 2W_a \mathrm{d}u\mathrm{d}x^a + \delta_{ab}\mathrm{d}x^a\mathrm{d}x^b\right|^{\frac{1+m}{2}} (\mathrm{d}u)^{-m}, \tag{6.54}$$

with $m \neq 1$, denote $x^1 = x$ and $x^2 = y$ and set

$$\tilde{H}(u,x,y) = kx^4 + \tfrac{1}{12}y^4, \qquad k = \frac{-1+3m-m^2}{12(m-1)^2}, \tag{6.55}$$

$$\rho(u,x,y) = x, \qquad W_1(u,x,y) = 0, \qquad W_2(u,x,y) = xy. \tag{6.56}$$

Then $\beta = \mathrm{d}u$ is null and it follows from Theorem 6.2.3 that F is not locally metrizable, hence F satisfies condition (i). On the other hand, it follows from the results in [44] that F is Ricci-flat. This shows that Ricci-flat m-Kropina spaces satisfying (i) do exist.

Part III

Vacuum Solutions in Finsler Gravity

Introduction to Part III: Vacuum Solutions in Finsler Gravity

In Part II we already briefly touched upon some aspects of the Finsler gravity field equation when we (locally) classified all Ricci-flat metrizable m-Kropina metrics. Such metrics are automatically vacuum solutions. Part III focuses completely on the field equation—which we properly introduce in Chapter 7—and exact vacuum solutions to it.

Most of the solutions that we present and investigate are of Berwald type. These are the topic of Chapter 8. Here, we first introduce a large class of general (α, β)-type solutions; we prove that when α is chosen to be a classical pp-wave metric and β is its defining covariantly constant null 1-form, then any (α, β)-metric constructed from these two building blocks will be an exact vacuum solution. In the specific case of Randers metrics of Berwald type, we also show that *any* Berwald vacuum solution must be of this type, resulting in a (local) classification of Berwald-Randers vacuum solutions. Completely analogous results hold for so-called *modified* Randers metrics, obtained as a small modification of the standard Randers metric and introduced because of their preferable causal properties.

An important question that arises is how such Finslerian spacetimes should be interpreted physically, and in particular, whether and how they can be physically distinguished from their general relativistic counterparts, given by the pseudo-Riemannian metric α. To answer this question we apply, in Chapter 9, a two-fold linearization scheme to the (α, β)-metric solutions obtained in Chapter 8 and conclude that the linearized solutions may be interpreted as Finslerian gravitational waves. We then study the observational signature of these grav-

itational waves by investigating the question of what would be observed in an interferometer experiment when such a Finslerian gravitational wave would pass the earth, and what would be the difference with the effect of a classical general relativistic gravitational wave. To this end, we compute the Finslerian radar distance, the typical observable measured in interferometer experiments. Remarkably, when interpreted correctly, the expression for the Finslerian radar distance turns out to be completely equivalent to its general relativistic analog, obtained in [45]. In other words, we show that gravitational wave interferometers are not able to distinguish between our Finsler gravitational waves and standard general relativistic ones, which is a remarkable conclusion.

Finally, in Chapter 10 we present a class of exact vacuum solutions of unicorn type. A particularly interesting subfamily of this class has cosmological symmetry, i.e. is spatially homogeneous and isotropic, and is additionally conformally flat, with conformal factor depending only on the timelike coordinate. We show that, just as in classical Friedmann-Lemaître-Robertson-Walker (FLRW) cosmology, this conformal factor may be interpreted as the scale factor of the universe. We compute this scale factor as a function of cosmological time, and we show that the solution corresponds to a linearly expanding (or contracting) Finslerian universe.

Since Part III is more physics-oriented than the preceding parts and generally deals with Finsler spaces of Lorentzian signature, we will (unless otherwise specified) assume that $\dim M = 4$ and use the convention that Greek indices $\mu, \nu, \rho, \sigma, \ldots$ run from 0 to $3 = \dim M - 1$, while Latin ones $i, j, k, l, \ldots$ run from 1 to 3. As in Part II, we will also encounter Latin indices $a, b, c \ldots$ from the beginning of the alphabet and these will run from 1 to $2 = \dim M - 2$, consistent with our earlier conventions.

CHAPTER 7

Finsler Gravity

The whole field of Finsler gravity can be concisely summarized by the statement that *Finsler gravity is just general relativity without the quadratic restriction*[1]. The 'quadratic restriction' refers to the fact that general relativity (GR) is based on pseudo-Riemannian geometry, in which the squared line-element $\mathrm{d}s^2$ is, by definition, quadratic in the coordinate displacements. The essential idea is that Finsler gravity generalizes GR by lifting this quadratic restriction—leading to Finsler geometry as the relevant underlying mathematical framework—while retaining other fundamental aspects of the theory as much as possible. As an example, the time measured by a clock between two events is still given, as it is in GR, by the length of the clock's spacetime trajectory connecting these events; in this case the Finslerian length rather than the pseudo-Riemannian length. In fact, Finsler geometry is the most general geometric framework that is compatible with this formulation of the clock postulate[2].

As discussed in some detail in the introduction there are various compelling arguments that motivate the study of Finsler gravity. While we will not repeat that discussion in detail here, we recall that much of this motivation comes from

(i) quantum gravity phenomenology [1] and specifically the idea of Lorentz in-

[1] This sentence is an adaptation of Chern's famous phrase saying that 'Finsler geometry is just Riemannian geometry without the quadratic restriction.'

[2] We remark that Weyl geometry, another generalization of Lorentzian geometry, is also compatible with the clock postulate, but in that case, the definition of proper time has to be revised [21].

variance violation (or deformed Lorentz symmetry), which has been linked to Finsler geometry [8, 9, 10, 11, 12, 13];

(ii) axiomatic approaches to spacetime and relativity [14, 15, 16, 17, 18, 19, 20, 87] that have been found to be compatible with geometric frameworks more general than pseudo-Riemannian geometry. And by the clock postulate argument given above, it seems that Finsler geometry is the most general framework one can aim for.

There are also hints that Finsler gravity could provide a more accurate description of the gravitational field of a kinetic gas, by taking into account the individual motion of each gas particle [88, 89], which is usually averaged over in the standard Einstein-Vlasov treatment [90, 91]. And there are hints that it might play a role in the understanding of dark matter and dark energy [92]. Furthermore, some teleparallel gravity models can be described within the framework of Finsler geometry as well [93].

Finally, apart from all of this, the simple fact that a theory as beautiful and successful as GR *can* be generalized in a natural geometric way, without breaking it, elicits a natural curiosity as to what such an extension may bring us. I would even argue that this on its own provides ample motivation for taking the idea seriously.

7.1 Finsler spacetimes

7.1.1 The basic mathematical definition

As the starting point for our mathematical definition of a Finsler spacetime, we will simply stick to Definition 1.2.3 with the additional requirement that the signature of $g_{\mu\nu}$ be Lorentzian in some subset of the domain $\mathcal{A}$ of the Finsler metric. We purposely remain vague about this subset for the moment but we will come back to it later. The homogeneity condition (1.6) ensures that the length is invariant under (orientation-preserving) parameterization and hence the time measured by a clock between two events can consistently be defined as the Finslerian length of the clock's spacetime trajectory connecting these events[3].

As a definition of a Finsler spacetime, this is a relatively weak one in the sense that there are many other definitions appearing in the literature that are more restrictive. In fact, various nonequivalent definitions exist (see e.g. [25, 26, 94,

[3]We remark that time-orientation needs to be taken into account as well in case the Finsler function is not symmetric, i.e. if $F(x, y) \neq F(x, -y)$.

28, 95, 96, 97]) and a consensus on what is the 'best' definition—if such a thing exists at all—has yet to be reached. It is not our current aim to significantly advance this endeavor, so the definition employed here has been chosen so as to be as general as possible. Indeed, most of the results we will discuss in what follows can be proven without further restrictions. It should be understood, however, that in order to guarantee that a viable physical interpretation is possible, the geometry should be subjected to more stringent requirements. Some of these are outlined below, but we will mostly discuss such considerations on a case-by-case basis throughout the chapters that follow.

7.1.2 Some remarks on causal structure and signature

Given a Finsler spacetime geometry, it is natural to postulate, in analogy with GR, that matter travels along timelike geodesics and light travels on null geodesics. The notion of a *null* direction is generalized in a mathematically straightforward way: a vector y^u at a point x^μ is said to be null (or lightlike) if $L(x, y) = g_{\mu\nu}(x, y)y^\mu y^\nu = 0$. However, the structure of the light cone, composed of such null vectors, may be nontrivial. In GR it is always the case that the light cone separates the tangent space at each point into three connected components, which we may interpret as consisting of forward-pointing timelike vectors, backward-pointing timelike vectors, and spacelike vectors, respectively. It is then consistent to define a timelike vector as one that has negative (or positive, depending on the sign convention) norm; that precisely singles out what we would intuitively call the inside or the interior[4] of the light cone, sometimes called the timelike cone. For a generic Finsler spacetime metric, however, these properties of the light cone structure are by no means guaranteed, and as such it is not obvious in general how to even define what one means by timelike vectors. It certainly is not obvious that these should be defined as vectors with negative norm. We will not consider this issue any further in its full generality here but rather discuss the details on a case-by-case basis. In many of the explicit examples of Finsler spacetimes that we will encounter, the causal structure (i.e. light cone structure) has exactly the desirable properties mentioned above, just as in the case of GR, allowing for a straightforward physical interpretation.

In a scenario where a natural (forward and/or backward) timelike cone can be identified, this cone should ideally be contained in the subbundle $\mathcal{A}$. Intuitively, this condition essentially states that the geometry is well defined for all timelike infinitesimal spacetime separations. Moreover, in the ideal case, the signature is Lorentzian throughout the entire timelike cone. That is, $g_{\mu\nu}(x, y)$ has Lorentzian

[4] We're not talking about the topological interior here.

signature whenever y is a timelike vector at x. It can be argued that it is not strictly necessary for the signature of $g_{\mu\nu}$ to be Lorentzian at spacelike vectors y, as it would not be possible, even in principle, to perform any physical experiment that probes such spacelike directions. This is true even in GR, but in that case $g_{\mu\nu}$ does not depend on y and so if the metric is Lorentzian in any direction it is automatically Lorentzian in all directions. In Finsler geometry this is not the case and hence it makes sense to make this distinction. For further arguments along these lines, we refer to [19] and references therein. Whether the light cone should be contained in $\mathcal{A}$ (or its topological closure) and whether the signature should be Lorentzian at the light cone is a more delicate question, which we will not further explore here. Mathematically such a property is very convenient (see e.g. [97, 40]), but many examples of interest simply do not seem to have this property.

7.2 The field equations in vacuum

We now turn to the field equation in Finsler gravity, which generalizes Einstein's field equation to the more general setting of Finsler spacetimes. While several earlier proposals exist in the literature [29, 30, 31, 98, 32, 33, 34, 35, 36], we consider here Pfeifer and Wohlfarth's field equation [28, 37, 38], which we deem to have the firmest basis.

7.2.1 Motivation for the field equations

Pfeifer and Wohlfarth's field equation can be derived, or rather motivated, by the following simple argument. In Newtonian gravity, the separation vector $\vec{v}$ between two nearby freely falling particles satisfies, in Cartesian coordinates, the deviation equation

$$\frac{\mathrm{d}^2 v^k}{\mathrm{d}t^2} + H^k{}_i v^i = 0, \qquad H^k{}_i = \delta^{kj} \frac{\partial^2 \phi}{\partial x^i \partial x^j}, \tag{7.1}$$

where $\phi(x)$ is the gravitational potential. In vacuum, the equation that determines the gravitational potential is the Laplace equation $\Delta\phi = \delta^{ij} \frac{\partial^2 \phi}{\partial x^i \partial x^j} = 0$, which can be written simply as $H^i{}_i = 0$, i.e. the *vanishing of the trace of the deviation tensor*. Of course, this principle extends to GR, where the gravitational field equation in vacuum is again equivalent to the vanishing of the trace of the geodesic deviation tensor, $R_{\mu\nu} = 0$.

We can immediately generalize this to the Finsler geometric setting. The

analogous situation is then governed by the Finslerian geodesic deviation equation (3.33) along a geodesic γ,

$$\frac{D^2 v^\rho}{Ds^2} + R^\rho{}_\mu(\gamma, \dot{\gamma}) v^\mu = 0, \tag{7.2}$$

where the geodesic deviation operator is given by $R^\rho{}_\mu(x, y) = R^\rho{}_{\mu\nu}(x, y) y^\nu$ in terms of the nonlinear curvature $R^\rho{}_{\mu\nu}$ of the Finsler metric. In analogy with Newtonian gravity and GR, it is natural to postulate that in the Finslerian case, the trace of the deviation tensor must vanish as well, leading to the candidate field equation $R^\mu{}_\mu = 0$. Since $R^\mu{}_\mu = \mathrm{Ric}$, the Finsler-Ricci scalar (see Section 2.3.3), this leads to the equation $\mathrm{Ric} = 0$ that was proposed by Rutz [27]. When restricted to Lorentzian metrics, Rutz's equation reduces to Einstein's vacuum equation $R_{\mu\nu} = 0$, since in that case $\mathrm{Ric} = R_{\mu\nu}(x) y^\mu y^\nu$. Hence the equation yields the correct general relativistic limit.

It turns out, however, that Rutz's equation is not variational; there does not exist an action functional that yields the Rutz equation as its Euler-Lagrange equation. Although this may not necessarily be problematic, it is not obvious in such a case how to couple the theory to matter, let alone how to eventually quantize the theory. The variational completion of Rutz's equation, i.e. the unique variational equation that is *as close to it as possible*, in a well defined sense [39], turns out to be the field equation that was proposed by Pfeifer and Wohlfarth in [28] using a Finsler extension of the Einstein-Hilbert action [38] (details below).

This is completely analogous to the situation in GR, where the correct[5] form of the vacuum field equation $R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0$ is also precisely the variational completion of Einstein's earlier proposal for a vacuum equation, $R_{\mu\nu} = 0$, obtained from the geodesic deviation argument above[6] [39]. While in the GR case, the variationally completed equation happens to be equivalent to the original one, this is not true any longer in the Finsler setting.

[5] Correct in the sense that it also leads to the correct matter coupling when the energy-momentum tensor is inserted on the RHS.

[6] We remark that Einstein himself arrived at both his early proposal $R_{\mu\nu} = 0$ as well as his final expression $R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0$ for the vacuum equation by alternative means [99].

7.2.2 Equivalent formulations of the field equation

In four spacetime dimensions, Pfeifer and Wohlfarth's field equation in vacuum reads[7] [28, 38]

$$-2\mathrm{Ric} + \frac{2L}{3} g^{\mu\nu} R_{\mu\nu} + \frac{2L}{3} g^{\mu\nu} \left(\bar{\partial}_\mu \left(\boldsymbol{\nabla}_D S_\nu \right) - S_\mu S_\nu + \nabla^C_{\delta_\mu} S_\nu \right) = 0, \tag{7.3}$$

in terms of (the horizontal part of) the Chern-Rund connection $\nabla^C_{\delta_\mu}$ and the dynamical covariant derivative $\boldsymbol{\nabla}_D$. This is essentially the Euler-Lagrange equation in four dimensions corresponding to the action functional

$$S = \int_{SM} \mathrm{Ric}\, \mathrm{d}\mu_{SM}, \tag{7.4}$$

where $SM = \{(x, \dot{x}) \in TM \,:\, F(x, \dot{x}) = 1\}$ is the unit tangent bundle, or *indicatrix*, with volume element $\mathrm{d}\mu_{SM}$. In order to make rigorous sense of this action functional and its critical points, however, one needs to use a construction in terms of compact subsets of the (positively) projective tangent bundle. When F is pseudo-Riemannian, (7.4) reduces to the Einstein-Hilbert action, (7.10) below, up to a boundary term. For details, we refer to [38].

Equivalently, the equation (7.3) can be expressed in terms of the Berwald connection rather than the Chern-Rund connection. Using the fact that the difference between the horizontal part of the Berwald and Chern-Rund connection is given by the Landsberg tensor (see Section 4.2), the field equation (7.3) can be rewritten as[8]

$$-2\mathrm{Ric} + \frac{2L}{3} g^{\mu\nu} R_{\mu\nu} + \frac{2L}{3} g^{\mu\nu} \left(\bar{\partial}_\mu \left(\boldsymbol{\nabla}_D S_\nu \right) + \nabla^B_{\delta_\mu} S_\nu \right) = 0. \tag{7.5}$$

This formulation[9] appears for instance in [44].

It is worth pointing out that a very similar equation has been derived in the case of positive definite Finsler metrics. In [100], Chen and Shen consider the action functional

$$S = \frac{1}{\mathrm{Vol}(SM)^{1-2/n}} \int_{SM} \mathrm{Ric}\, \mathrm{d}\mu_{SM}, \tag{7.6}$$

[7] We note that in [28] the Landsberg tensor is defined with a relative minus sign with respect to our definition and in both references [28, 38] the Ricci curvatures are defined with a relative minus sign with respect to our definition.

[8] Since $g^{ij} \nabla^B_{\delta_i} S_j = g^{ij} \left(\nabla^C_{\delta_i} S_j - S_i S_j \right)$.

[9] We note that also in this reference both the Landsberg tensor and the Ricci tensor(s) are defined with a relative minus sign with respect to our definitions.

where $n = \dim M$. This is the normalized version of (7.4). In terms of the Chern-Rund connection, the Euler-Lagrange equations in four dimensions read

$$-2\mathrm{Ric} + \frac{2L}{3} g^{\mu\nu} R_{\mu\nu} + \frac{2L}{3} g^{\mu\nu} \left(\bar{\partial}_\mu \left(\boldsymbol{\nabla}_D S_\nu \right) - S_\mu S_\nu + \nabla^C_{\delta_\mu} S_\nu \right) = \frac{2L}{3} r, \tag{7.7}$$

where

$$r = \frac{1}{\mathrm{Vol}(SM)} \int_{SM} \mathrm{Ric}\, \mathrm{d}\mu_{SM} \tag{7.8}$$

is the average value of Ric on SM. Clearly the only difference between (7.3) and (7.7) is the presence of a possibly nonvanishing (constant) value of r on the RHS, playing the role of a kind of cosmological constant. Indeed, in the case that the geometry is pseudo-Riemannian, (7.7) reduces to the equation for an Einstein space,

$$R_{\mu\nu} = r g_{\mu\nu}, \tag{7.9}$$

or, in other words, the vacuum field equation for GR in the presence of a possibly nonvanishing cosmological constant $\Lambda = r$. Similarly, the Pfeifer-Wohlfarth equation (7.3) reduces to Einstein's equation in vacuum without cosmological constant, $R_{\mu\nu} = 0$.

For reference, it is worth briefly comparing the results discussed above with the classical results for Riemannian metrics. It is well-known (see e.g. [101]) that for Riemannian metrics on a compact manifold M the critical points of the Einstein-Hilbert action functional

$$S_{\mathrm{EH}} = \int_M R\, \mathrm{d}\mu_M, \tag{7.10}$$

are precisely the Ricci-flat metrics, $R_{\mu\nu} = 0$, while the critical points of the normalized Einstein-Hilbert action

$$S = \frac{1}{\mathrm{Vol}(M)^{1-2/n}} \int_M R\, \mathrm{d}\mu_M, \tag{7.11}$$

are precisely the Einstein metrics, $R_{\mu\nu} = \Lambda g_{\mu\nu}$.

We will refer to Pfeifer and Wohlfarth's field equation, (7.3) simply as the field equation in Finsler gravity, and solutions to it will be referred to simply as solutions in Finsler gravity.

7.2.3 Weakly Landsberg and Berwald field equation

The field equations simplify considerably if one restricts attention to the class of weakly Landsberg metrics (see Section 4.2), defined by the condition that $S_\mu = 0$, which holds in particular for all Berwald spaces as well as all Landsberg spaces. In that case, all terms involving the mean Landsberg curvature vanish identically and (7.3) reduces to

$$\mathrm{Ric} - \frac{L}{3} g^{\mu\nu} R_{\mu\nu} = 0, \tag{7.12}$$

or equivalently,

$$\left(L g^{\mu\nu} - 3 y^\mu y^\nu\right) R_{\mu\nu} = 0. \tag{7.13}$$

From this, it follows that within the class of weakly Landsberg spaces, the vanishing of the Finsler-Ricci tensor, $R_{\mu\nu} = 0$, is a sufficient condition for being a vacuum solution to the field equations. It turns out that in many cases of interest (but not in general) it is a necessary condition as well. For instance, for Randers metrics of Berwald type, the vacuum field equation is equivalent to $R_{\mu\nu} = 0$, as we will prove in Section 8.2, and similar results have been obtained in the case of null m-Kropina metrics [44] and Finsler spacetimes satisfying strict smoothness requirements [40]. This equivalence also holds for spherically symmetric and asymptotically flat Berwald spacetimes; details will be provided in a forthcoming article. By a recent generalization of Birkhoff's theorem to spherically symmetric, Ricci-flat Berwald spaces [102], this implies in particular, that any vacuum solution of Berwald type that is spherically symmetric and asymptotically flat is either flat (in the sense that the nonlinear curvature tensor vanishes identically) or given by the Schwarzschild metric. In general, the vacuum equation is not equivalent to Ricci-flatness, though: non–Ricci-flat vacuum solutions do exist. An explicit example illustrating this will also be provided in a forthcoming article.

Since any Berwald space is weakly Landsberg, (7.13) applies in particular to Berwald spaces. In this case $R_{\mu\nu} = R_{\mu\nu}(x)$ is independent of y, which is an incredible convenience, computationally. When L is pseudo-Riemannian, and hence, in particular, Berwald, the inverse fundamental tensor $g^{\mu\nu} = g^{\mu\nu}(x)$ is also independent of y and hence both terms in the field equation (7.13) are quadratic in y. Differentiating twice with respect to the y-coordinates implies that $3R_{\mu\nu} - g_{\mu\nu}R = 0$ and taking the trace then reveals that $R = 0$ and hence $R_{\mu\nu} = 0$. Hence for pseudo-Riemannian metrics, the field equation in vacuum indeed reduces to Einstein's field equations in vacuum.

7.2.4 Exact solutions

The list of known exact solutions in Finsler gravity is very short. Indeed, prior to the contributions on which this dissertation is based, the only[10] ones known in the literature were the

(i) (m-Kropina type) Finsler pp-waves [43] and their generalizations as

(ii) *very general relativity* (VGR) spacetimes [44].

Over the last couple of years, this list has been extended by the

(iii) Randers pp-waves [H4] with their extension to

(iv) generic (α, β)-metric pp-waves [H5], and finally

(v) the unicorn solutions [H6].

Additionally, vacuum solutions of Berwald type with cosmological symmetry have been classified and properly Finslerian spherically symmetric vacuum solutions have been obtained, the details of which will appear in future publications. It is the aim of the chapters that follow to properly introduce and discuss the solutions (*iii*)–(*v*) in detail.

[10] Specifically in the realm of positive definite Finsler geometry, an additional class of solutions has been obtained [100].

CHAPTER 8

Exact Vacuum Solutions of Berwald Type

In this chapter, we present and investigate those of our solutions that are of Berwald type. We begin in Section 8.1 with generic (α, β)-metric solutions, then specialize to Randers metrics in Section 8.2, and to modified Randers metrics in Section 8.3. Regarding (α, β)-metrics, we stick to the notation introduced in Section 5.2.1 and we recall, in particular, that by some abuse of language, we will refer to either of α, A, a and a_{ij} as the pseudo-Riemannian metric, and to β, b and b_i as the 1-form, and sometimes we will write expressions such as $\beta = \mathrm{d}t$ and $\alpha = \sqrt{|-\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2|}$, but the interpretation of such expressions should be clear. In the presence of an (α, β)-metric, indices will be raised and lowered with a_{ij} unless otherwise specified. Moreover, when discussing properties of the 1-form β such as being 'covariantly constant' (or 'parallel', which has the same meaning) or 'null', we will always mean that β is covariantly constant *with respect to* α or null *with respect to* α, respectively, unless otherwise specified.

8.1 (α, β)-metrics

The results in this section are based mostly on [H5].

8.1.1 Construction of exact solutions

From a physical perspective, (α, β)-metrics (see Section 5.2.1) allow us to deform a general relativistic (i.e. Lorentzian) spacetime α into a Finsler spacetime,

given a 1-form β. And it turns out, as we will prove below, that these types of metrics can be used to generalize some of the vacuum solutions to Einstein's field equations to properly Finslerian vacuum solutions to Pfeifer and Wohlfarth's field equation for Finsler gravity. This procedure is possible whenever such a classical solution admits a covariantly constant vector field, or equivalently, 1-form. More precisely, if the Lorentzian metric α solves the classical Einstein equations in vacuum and the 1-form β is covariantly constant with respect to α then any (α, β)-metric constructed from the given α and β is a solution to the Finslerian field equations.

To see why this is true, we first recall Corollary 5.2.2, which states that if β is covariantly constant with respect to α then any (α, β)-metric F constructed from these building blocks is of Berwald type and the affine connection of F coincides with the (Levi-Civita) affine connection of α.

If the affine connections coincide, the associated curvature tensors and in particular the (affine) Ricci tensors must also coincide. So if α happens to be a vacuum solution to Einstein gravity, i.e. its Ricci tensor vanishes, then it follows that the affine Ricci tensor of F must vanish as well. And since the Finsler-Ricci tensor is the symmetrization of the affine Ricci tensor, by Lemma 4.1.1, the Finsler-Ricci tensor vanishes as well. This implies that (7.13) is satisfied, i.e. F is a vacuum solution in Finsler gravity. We may summarize this result with the following theorem.

Theorem 8.1.1. *Let F be any (α, β)-metric such that:*

(i) α solves the classical Einstein equations in vacuum;

(ii) β is covariantly constant.

Then F is a Ricci-flat vacuum solution to the field equation in Finsler gravity. Moreover, the affine connection of F coincides with the Levi-Civita connection of α.

In this way, (α, β)-metrics provide a mechanism to *Finslerize* any vacuum solution to Einstein's field equations, as long as the solution admits a covariantly 1-form or equivalently, a covariantly constant vector field. The theorem generalizes some of the results obtained in [H4] for Randers metrics and in [43, 44] for m-Kropina metrics (or VGR spacetimes) to arbitrary Finsler spacetimes with (α, β)-metric. In particular, all pp-wave type solutions in Finsler gravity currently known in the literature are of this type.

8.1.2 pp-Wave solutions

Let us investigate the solutions of the type of Theorem 8.1.1 in some more detail. It turns out that if a vacuum solution α to Einstein's field equations admits a covariantly constant 1-form β then at least one of the following conditions must hold:

1. α is flat;
2. β is null.

This was shown for the first time in [103] (see also [104, 105]) and in Appendix A.2 we provide a self-contained proof, starting from the Frobenius theorem and culminating in Corollary A.2.3. We remark that the result assumes that the dimension of the pseudo-Riemannian manifold is 4; it is generally not true in higher dimensions. Hence Theorem 8.1.1 leads to two classes of solutions.

Class 1: flat solutions

If α (and hence F) is flat then in coordinates in which all Christoffel symbols vanish, β must have constant components. Hence the resulting class of solutions is in some sense trivial—yet not necessarily uninteresting. Provided α is of Lorentzian signature[1], it can always be written in suitable coordinates in the following way, where we stick to the notation introduced in Section 5.2.1.

Class 1: flat solutions

Any (α, β)-metric constructed from

$$a = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2, \qquad b = b_\mu \mathrm{d}x^\mu, \tag{8.1}$$

where b_μ = constant, is a vacuum solution to the field equations in Finsler gravity. The resulting geometry is of Berwald type with all affine connection coefficients vanishing identically in these coordinates.

Class 2: pp-wave solutions

The second possibility, the one where β is null, leads to a richer and more interesting class of solutions. In this case, α is a CCNV metric, meaning that it admits a covariantly constant null vector (CCNV), namely β, or rather its vector equivalent via the isomorphism induced by α. In Lorentzian signature,

[1] The result can easily be generalized to other signatures as well.

CCNV metrics are also known as *pp-waves* (plane-fronted gravitational waves with parallel rays) and these have been studied in detail in [103, 86] (see [79, Section 24.5] for a summary).

In Part II we have already encountered such metrics. In particular, we proved the important Lemma 6.4.5. Combined with Proposition A.2.1, and using the fact that α is a vacuum solution and hence Ricci-flat, this lemma says that coordinates (u, v, x, y) can always be chosen such that α and β attain the form

$$a = -2\mathrm{d}u\mathrm{d}v + H(u, x, y)\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2, \qquad b = \mathrm{d}u, \tag{8.2}$$

and such that $\Delta_{(x,y)}H = (\partial_x^2 + \partial_y^2)H = 0$. The last statement also follows easily from the fact that the Ricci tensor of the metric a is given (up to a nonzero constant) by the (x, y)-Laplacian $\Delta_{(x,y)}H$ of H. We may therefore characterize the second class of solutions in the following way—again assuming for concreteness that α is a Lorentzian metric.

Class 2: pp-wave solutions

Any (α, β)-metric constructed from

$$a = -2\mathrm{d}u\mathrm{d}v + H(u, x, y)\,\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2, \tag{8.3}$$

$$b = \mathrm{d}u, \tag{8.4}$$

such that $\Delta_{(x,y)}H = 0$, is a vacuum solution to the field equations in Finsler gravity. The resulting geometry is of Berwald type with affine connection identical to the Levi-Civita connection of a.

Note that when $H = 0$ the geometries in *Class 2* are also contained in *Class 1*. It is not the case, however, that *Class 1* is a subset of *Class 2* because in *Class 1* the 1-form β need not be null, necessarily. The preceding line of argument shows that these two classes of solutions in fact exhaust all possibilities, which we encapsulate in the following theorem.

Theorem 8.1.2. *Any (α, β)-metric such that*

(i) α solves the classical Einstein equations in vacuum;

(ii) β is covariantly constant,

(i.e. any vacuum solution of the type of Theorem 8.1.1) must belong to either Class 1 or Class 2, provided that α has Lorentzian signature.

Before we move on to solutions of plane wave type, we end this section with some remarks about the relation between our (α, β)-type solutions and other solutions in the literature.

- The (m-Kropina type) Finsler pp-waves [43] and the Randers pp-waves obtained in [H4] are special cases of the solutions introduced here.
- For Randers metrics of Berwald type *any* vacuum solution must be of the type described in Theorem 8.1.1, that is, α is necessarily a vacuum solution in Einstein gravity and β is necessarily covariantly constant. Any such solution is therefore either in *Class 1* or *Class 2*. This was first shown in [H4] and we will prove it in Section 8.2.
- For m-Kropina metrics, some of the vacuum solutions obtained [44] are of a more general type than the ones introduced here, as illustrated by Example 6.4.7; our solutions are all trivially metrizable (by α), whereas the metric in Example 6.4.7 is not metrizable. In fact, our results from Section 6.4 show that any of the solutions obtained in [44] belongs to either *Class 1* or *Class 2* if and only if the geometry is metrizable, and Example 6.2.4 provides the precise necessary and sufficient condition for this to be the case.

This list comprises all exact solutions to Pfeifer and Wohlfarth's equation for Finsler gravity currently known in the literature, apart from:

- the unicorn solutions [H6] that are the topic of Chapter 10;
- pseudo-Riemannian metrics that solve Einstein's vacuum field equation, which are trivially vacuum solutions in Finsler gravity as well.

8.1.3 Plane wave solutions in Brinkman and Rosen coordinates

Equation (8.3) expresses the pp-wave metric in Brinkmann form [106]. In this form the metric reads

$$a = -2\mathrm{d}u\mathrm{d}v + H(u, x^a)\,\mathrm{d}u^2 + \delta_{ab}\,\mathrm{d}x^a\mathrm{d}x^b, \tag{8.5}$$

where we have used the notation $x = x^1$, $y = y^1$, which will be somewhat more convenient for our present purposes, and where indices $a, b, c \ldots$ run from 1 to 2, consistent with the conventions we have stuck to all along. For the description of the physical effects of (plane) gravitational waves in general relativity, it is sometimes more convenient to use a different coordinate system, known as

Rosen coordinates [107]. In our experience, this remains true in the Finsler setting. When we compute the effect on the radar distance of a passing (linearized) Randers gravitational wave in Chapter 9, our starting point will be the expression for the (exact) wave in Rosen coordinates. Therefore we briefly review the relation between the two coordinate systems here.

Rosen coordinates can be introduced for the subclass of pp-waves known as *plane waves*. These are characterized by the property that the curvature tensor does not change (i.e. is covariantly constant) along the Euclidean 'wave surfaces' given in Brinkmann coordinates by $\mathrm{d}u = \mathrm{d}v = 0$, i.e.

$$\nabla_{\partial_{x^1}} R^{\rho}{}_{\sigma\mu\nu} = \nabla_{\partial_{x^2}} R^{\rho}{}_{\sigma\mu\nu} = 0. \tag{8.6}$$

We remark as a side note that $\nabla_{\partial_v} R^{\rho}{}_{\sigma\mu\nu} = 0$ always holds, identically, so invariance along $\mathrm{d}u = \mathrm{d}v = 0$ is actually equivalent to invariance along the hypersurfaces $\mathrm{d}u = 0$. The conditions (8.6) are equivalent to the statement that $\partial_a\partial_b\partial_c H = 0$ in Brinkmann coordinates (8.5), i.e. that $H(u, x, y)$ is a quadratic polynomial in x^a (for each fixed value of u). In that case there always exists a coordinate transformation that removes the linear and constant terms [79, Section 24.5] and brings the metric into the form

$$a = -2\mathrm{d}u\mathrm{d}v + A_{ab}(u)x^a x^b\,\mathrm{d}u^2 + \delta_{ab}\,\mathrm{d}x^a\mathrm{d}x^b \tag{8.7}$$

This is the standard expression for a plane-wave metric in Brinkmann form. Moreover, an argument completely analogous to the one given in the proof of Lemma 6.4.5 shows that we may assume without loss of generality that the 1-form $b = \mathrm{d}u$ remains unchanged under the transformation to these new coordinates.

On the other hand, any such plane wave metric can also be written in Rosen form

$$a = -2\mathrm{d}U\mathrm{d}V + h_{ij}(U)\mathrm{d}y^i\mathrm{d}y^j, \tag{8.8}$$

where h_{ij} is a two-dimensional Riemannian metric. And conversely, any metric of Rosen type (8.8) can be cast in the form (8.7). The two coordinate systems are related via

$$U = u, \quad V = v - \frac{1}{2}\dot{E}_{ai}E^i{}_b x^a x^b, \quad x^a = E^a{}_i y^i, \tag{8.9}$$

where $A_{ab} = \ddot{E}_{ai}E^i{}_b$ and $E^a{}_i$ is a vielbein for h_{ij} in the sense that $h_{ij} = E^a{}_i E^b{}_j \delta_{ab}$, satisfying the additional symmetry condition $\dot{E}_{ai}E^i{}_b = \dot{E}_{bi}E^i{}_a$. Such a vielbein can always be chosen. For details, we recommend the lecture notes

[108] and references therein (see also the Appendix of [109]). Note that we have momentarily labeled the y-coordinates by indices $i, j, k, \ldots$ so as to distinguish them from indices $a, b, c, \ldots$ in order that we may employ the usual notation with regard to the vielbein indices: $E^i{}_a$ represents the (matrix) inverse of $E^a{}_i$ and indices $a, b, c \ldots$ are raised and lowered with δ_{ab}, whereas indices $i, j, k, \ldots$ are raised and lowered with h_{ij}. In this case $i, j, k, \ldots$ thus take the values 1 and 2, contrary to our standard convention. The dot that sometimes appears above the vielbein represents a U-derivative. Since the vielbein depends only on $U = u$, this derivative is equivalent to a u-derivative, and moreover, the act of raising and lowering the $a, b, c, \ldots$ indices commutes with taking such a derivative of the vielbein.

It is again the case that, after relabeling $U, V \mapsto u, v$, the form of the 1-form $b = \mathrm{d}u = \mathrm{d}U$ remains unchanged under this transformation, which in this case follows trivially, since $u = U$. After also relabelling $y \mapsto x$, we conclude that we can express any *Class 2* solution of plane-wave type in Rosen coordinates as follows,

$$F = \alpha\, \phi(\beta/\alpha), \qquad a = -2\mathrm{d}u\mathrm{d}v + h_{ij}(u)\mathrm{d}x^i\mathrm{d}x^j, \qquad b = \mathrm{d}u, \tag{8.10}$$

Conversely, for any choice of ϕ and $h_{ij}(u)$, this is a vacuum solution to the field equations in Finsler gravity if a is a vacuum solution to Einstein's field equation. The resulting geometry is of Berwald type with affine connection identical to the Levi-Civita connection of a. The latter follows from the line of argument above but is also easy to check explicitly since all Christoffel symbols of a in Rosen coordinates with an upper index u vanish identically, and as a consequence, $b = \mathrm{d}u$ is covariantly constant (as we already knew, of course), so that the result follows by Theorem 8.1.1.

8.2 Randers metrics

Next, we turn to Randers metrics [75], i.e. (α, β)-metrics of the specific form $F = \alpha + \beta$. We already encountered such metrics in Chapter 5. Apart from their intrinsic value as relatively simple and computable, yet nontrivial examples of Finsler metrics, Randers metrics have also found a wealth of applications in a wide variety of fields such as biology, psychology, and physics. Most of the results developed in this section are based on [H4]. We point out, though, that [H4] uses a signature convention that is opposite to the one employed here. Before turning to the field equations, we briefly discuss the domain of definition of the standard Randers metric.

8.2.1 Domain of definition

In the case where α is positive definite, it is well-known (see e.g. [110]) that $F = \alpha + \beta$ is a positive definite Finsler metric if and only if $|b| < 1$. A crucial ingredient that leads to this conclusion is the Cauchy-Schwarz inequality $|\beta| \leq |b|\alpha$, which does not extend in this simple form to indefinite signatures. In the latter case, the situation therefore becomes somewhat more complex. It can be shown using the general formula (C.14) for the determinant of the fundamental tensor of an (α, β)-metric that

$$\det g = \left(\frac{\alpha + \beta}{\alpha}\right)^{n+1} \det a \tag{8.11}$$

for Randers metrics, where $n = \dim M$. This expression has in general no well-defined limit as $\alpha \to 0$, so from this, it is clear that in Lorentzian signature, the subbundle $\mathcal{A}$ can only include points (x, y) for which $\alpha \neq 0$. In order to have a connected subbundle, one might propose to work on the subbundle consisting of α-timelike vectors, i.e. the (forward and backward) timelike cone of α. However, this does not work in general; one has to restrict $\mathcal{A}$ just a little bit further. The following result was proven in [H4] (cf. Section 8.3 and [111] for related results).

Proposition 8.2.1. *Given a Lorentzian metric $\alpha = \sqrt{|a_{\mu\nu}y^\mu y^\nu|}$ on a manifold M with time-orientation T and a past-pointing 1-form β that is either null or timelike with respect to a, the Randers metric $F = \alpha + \beta$ defines a Finsler spacetime with Lorentzian signature on*

$$\mathcal{A} = \{(x, y) \in TM \,:\, a_{\mu\nu}(x)y^\mu y^\nu < 0,\ a_{\mu\nu}(x)T^\mu(x)y^\nu < 0\}\,, \tag{8.12}$$

i.e. the forward timelike cone of α.

In order to prove this, we start with some lemmas. Note that the result of Proposition 8.2.1, as well as that of Lemma 8.2.3, is dependent on the signature convention, which we recall has been chosen here as $(-,+,+,+)$. The first lemma is a standard result.

Lemma 8.2.2. *If $\{M_\lambda\}_{\lambda\in[0,1]}$ is a continuous family of $n \times n$-matrices such that $\det M_\lambda \neq 0$ for all λ, then the signature of M_λ is constant for $\lambda \in [0, 1]$.*

Proof. We use the well-known fact that the n eigenvalues of M_λ each depend continuously on λ (see e.g. [112, Cor. VI.1.6]). We need to show that the respective signs of these n eigenvalue functions are constant. Suppose for contradiction

that the k-th eigenvalue is positive for some λ_1 and negative for some λ_2. By continuity, it follows from the intermediate value theorem that there must exist some λ between λ_1 and λ_2 for which the eigenvalue vanishes. In that case, the determinant of M_λ vanishes for that value of λ, which is a contradiction. Hence the result follows. □

Lemma 8.2.3. *Let y be a timelike and future-oriented vector at $x \in M$, and b a timelike or null, and past-oriented vector at x, all with respect to some time-oriented Lorentzian metric $g_{\mu\nu}$. Then $g_{\mu\nu}y^\mu b^\nu > 0$.*

Proof. Let the time orientation be defined in terms of a nowhere-vanishing timelike vector field v. Then the future-orientation of y and past-orientation of b imply that $g_{\mu\nu}y^\mu v^\nu < 0$ and $g_{\mu\nu}b^\mu v^\nu > 0$. Now notice that we can always pick a basis of T_xM such that v has coordinates $(v^0, 0, \ldots, 0)$ with $v^0 > 0$ and such that $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric (first pick a basis such that the metric has the form of the standard Minkowski metric, then apply a suitable spatial rotation so that all spatial components of v^μ except v^1 are mapped to 0, and finally apply an appropriate one-dimensional Lorentz boost to make v^1 vanish as well). In this basis we write $\vec{y} = (y^1, y^2, \ldots, y^{n-1})$. With this notation, the future– and past-pointing properties of y and b, respectively, translate to $y^0 > 0$ and $b^0 < 0$ and the remaining properties can be stated as $|y^0| > |\vec{y}|$ and $|b^0| \geq |\vec{b}|$. Combined they read $|\vec{y}| < y^0$ and $|\vec{b}| \leq -b^0$. Then we have $|\vec{y}||\vec{b}| < -y^0b^0$ and hence, in terms of the standard dot product on $\mathbb{R}^{n-1}$, we have

$$\vec{y} \cdot \vec{b} \geq -|\vec{y}||\vec{b}| > y^0b^0, \tag{8.13}$$

by the Cauchy-Schwarz inequality, $-\vec{y} \cdot \vec{b} \leq |\vec{y} \cdot \vec{b}| \leq |\vec{y}||\vec{b}|$. It thus follows that

$$g_{\mu\nu}y^\mu b^\nu = \eta_{\mu\nu}y^\mu b^\nu = -y^0b^0 + \vec{y} \cdot \vec{b} > -y^0b^0 + y^0b^0 = 0. \tag{8.14}$$

□

Lemma 8.2.3 thus says that under the stated conditions, $\beta > 0$. The proof of Proposition 8.2.1 is now quite simple.

Proof of Proposition 8.2.1. Lemma 8.2.3 shows that $\beta > 0$ and hence that $F = \alpha + \beta > 0$ on the forward timelike cone of α, so that $\det g$ is strictly negative, by (8.11). The remainder of the proof is similar to its counterpart in the positive definite scenario. Introduce a parameter λ and consider the family of functions

$F_\lambda = \alpha + \lambda\beta$ for $\lambda \in [0,1]$. For each value of λ, $\det g_\lambda$ is strictly negative and so, by Lemma 8.2.2, g_λ has constant signature for $\lambda \in [0,1]$. In particular $g_{\lambda=1}$ has the same signature as $g_{\lambda=0}$, which is just the statement that g is Lorentzian. Since the homogeneity of F is clear, this completes the proof that F defines a Finsler spacetime on the forward timelike cone of α with Lorentzian signature on all of $\mathcal{A}$. □

8.2.2 The field equations

In general, as discussed, the vanishing of the Finsler-Ricci tensor is a sufficient condition for a Berwald spacetime to be a solution to Pfeifer and Wohlfarth's field equation (7.13). For Berwald-Randers metrics it is a sufficient condition as well.

Proposition 8.2.4. *For Berwald-Randers spacetimes, the field equation is equivalent to Ricci-flatness, i.e.*

$$(Lg^{\mu\nu} - 3y^\mu y^\nu)\, R_{\mu\nu} = 0 \quad \Leftrightarrow \quad R_{\mu\nu} = 0 \quad \Leftrightarrow \quad \mathrm{Ric} = 0.$$

Proof. The equivalence between the Ric $= 0$ and $R_{\mu\nu} = 0$ holds in general, and each of these trivially implies the field equation on the left. The non-trivial part of the proof consists of showing that the field equation implies Ricci-flatness. We first compute the fundamental tensor for a Randers spacetime $F = \alpha + \beta$. A straightforward computation using that $\partial_\mu \alpha = \epsilon y_\mu/\alpha$ and $\partial_\mu \partial_\nu \alpha = \frac{1}{\alpha}(\epsilon a_{\mu\nu} - y_\mu y_\nu/\alpha^2)$, where $\epsilon = \mathrm{sgn}(a_{\mu\nu} y^\mu y^\nu)$, shows that

$$g_{\mu\nu} = \frac{\epsilon F}{\alpha} a_{\mu\nu} + b_\mu b_\nu + (b_\mu y_\nu + b_\nu y_\mu)\frac{\epsilon}{\alpha} - y_\mu y_\nu \frac{\beta}{\alpha^3}. \tag{8.15}$$

Although a little tedious, it's also straightforward to check that its inverse is given by

$$g^{\mu\nu} = \frac{\epsilon\,\alpha}{F} a^{\mu\nu} - (b^\nu y^\mu + b^\mu y^\nu)\frac{\epsilon\,\alpha}{F^2} + y^\mu y^\nu \frac{(\beta + \epsilon|b|^2\alpha)}{F^3}\,. \tag{8.16}$$

Substituting this into the equation $(Lg^{\mu\nu} - 3y^\mu y^\nu)\, R_{\mu\nu} = 0$ and using the fact that $R_{\mu\nu}$ is symmetric leads to

$$\begin{aligned}\epsilon R\alpha(\alpha+\beta)^2 - 2\epsilon\alpha(\alpha+\beta)R_{\mu\nu}b^\mu y^\nu + (\beta + \epsilon|b|^2\alpha)R_{\mu\nu}y^\mu y^\nu \\ - 3(\alpha+\beta)R_{\mu\nu}y^\mu y^\nu = 0,\end{aligned} \tag{8.17}$$

where we have denoted $R \coloneqq a^{\mu\nu}R_{\mu\nu}$. This is a polynomial equation in α and we can separate it into a rational part (even powers of α) and an irrational part (odd powers of α) as

$$\epsilon R\alpha^3 + (\epsilon R\beta^2 - 2\epsilon\beta R_{\mu\nu}y^\mu b^\nu + (\epsilon|b|^2 - 3)R_{\mu\nu}y^\mu y^\nu)\alpha \\ = 2\epsilon(R_{\mu\nu}b^\mu y^\nu - R\beta)\alpha^2 + 2\beta R_{\mu\nu}y^\mu y^\nu \,. \tag{8.18}$$

On either of the open sets $\pm A > 0$ (i.e. $\epsilon = \pm 1 = \text{const}$), (8.18) is of the form $P\alpha = Q$, where P and Q are both polynomial in y. We claim that this implies that $P = Q = 0$. To see this, suppose for contradiction that $P \neq 0$. Then we would be able to write $\alpha = Q/P$ and hence α would be a rational function of y, which contradicts Lemma 5.3.2. Hence we must indeed have $P = 0$, which implies that $Q = 0$ as well. It thus follows that both sides of (8.18) must vanish separately. We focus on the vanishing of the rational part,

$$2(R_{\mu\nu}b^\mu y^\nu - R\beta)A + 2\beta R_{\mu\nu}y^\mu y^\nu = 0\,, \tag{8.19}$$

where we have used that $A = \epsilon\alpha^2$. Differentiating twice with respect to y (while keeping in mind that $R_{\mu\nu}$ depends only on x, by the Berwald property) and contracting with $a^{\mu\nu}$ leads to

$$4(8R_{\mu\nu}b^\mu y^\nu - 5R\beta) = 0, \tag{8.20}$$

and substituting this back into equation (8.19) yields

$$\beta\left(2R_{\mu\nu}y^\mu y^\nu - \tfrac{3}{4}RA\right) = 0, \tag{8.21}$$

which implies that the expression in parentheses must vanish. This is obviously true whenever $\beta \neq 0$, but the remaining points where $\beta = 0$ constitute a set with empty topological interior, and hence the conclusion that the expression in parentheses must vanish can be extended by continuity to this set as well. Now we perform the same trick once again; differentiating this expression twice with respect to y and contracting with $a^{\mu\nu}$ results in $-2R = 0$. Substituting this back into (8.21) yields $\mathrm{Ric} = R_{\mu\nu}y^\mu y^\nu = 0$, as desired. □

In other words, in the context of Berwald-Randers spacetimes, the vacuum field equations of Finsler gravity are formally identical to the vacuum field equations of general relativity and they also coincide with Rutz's equation [27], $\mathrm{Ric} = 0$. It turns out that these equivalences hold more generally under strict additional smoothness requirements on the Finsler Lagrangian [40]. Without such additional constraints, however, the equivalence does not extend even to general

(α, β)-metrics. This will be demonstrated in a forthcoming article.

Remark 8.2.5. *In the final stages of writing this dissertation it came to our attention (through discussions with Nicoleta Voicu) that the proof of Proposition 8.2.4 can be simplified by noticing the following. Since b^μ is necessarily covariantly constant for a Berwald-Randers metric by Proposition 5.3.4, it follows directly from the definition of the Riemann curvature tensor $\bar{R}^\rho{}_{\sigma\mu\nu}$ of $a_{\mu\nu}$ that $\bar{R}^\rho{}_{\sigma\mu\nu}b^\sigma = 0$, and hence the Ricci-tensor of $a_{\mu\nu}$ satisfies $\bar{R}_{\mu\nu}b^\nu = 0$. Since $\bar{R}_{\mu\nu} = R_{\mu\nu}$ it follows that all occurrences of $R_{\mu\nu}b^\nu$ in the proof may be replaced by zero immediately.*

8.2.3 Randers pp-waves

Proposition 8.2.4 allows us to classify all Berwald-Randers vacuum solutions to Pfeifer and Wohlfarth's field equation. The first step toward this end is to reduce the problem to a classification problem in general relativity rather than Finsler gravity.

Theorem 8.2.6. *A Randers metric of Berwald type, $F = \alpha + \beta$, is a vacuum solution in Finsler gravity if and only if α is a vacuum solution to Einstein's field equation.*

Proof. Since F is Berwald-Randers, Proposition 5.3.4 implies that its affine connection coincides with the Levi-Civita connection of α. It follows that the associated curvature tensors and in particular the (affine) Ricci tensors must also coincide. In particular, the affine Ricci tensor is symmetric and hence it is equal to its symmetrization, the Finsler-Ricci tensor. Hence the Finsler-Ricci tensor vanishes if and only if the Ricci tensor of α vanishes. By Proposition 8.2.4 this implies the result. □

Taking into account Proposition 5.3.4, it is clear that Theorem 8.2.6 can also be formulated in the following alternative way.

Theorem 8.2.7. *For a Randers metric $F = \alpha + \beta$, the following are equivalent:*

(i) F is a Berwald vacuum solution in Finsler gravity;

(ii) α solves Einstein's equation in vacuum and F is Berwald;

(iii) α solves Einstein's equation in vacuum and β is covariantly constant.

Hence any vacuum solution of Berwald-Randers type is of the type of Theorem 8.1.1. Combining this with Theorem 8.1.2 leads to the conclusion that any Berwald-Randers metric that solves the vacuum field equation in Finsler gravity must belong to either *Class 1* or *Class 2* in the terminology introduced in Section 8.1.2—provided α has Lorentzian signature. We thus arrive at the following (local) classification.

Corollary 8.2.8. *A Randers metric $F = \alpha + \beta$ of Berwald type, where α is a Lorentzian metric, is a vacuum solution in Finsler gravity if and only if one of the following statements is true:*

- *There exist local coordinates $(x^\mu) = (u, v, x, y)$ such that*

$$F = \alpha + \beta = \sqrt{|-2\mathrm{d}u\mathrm{d}v + H(u,x,y)\,\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2|} + \mathrm{d}u, \tag{8.22}$$

  with $\Delta_{(x,y)} H = 0$;

- *There exist local coordinates $(x^\mu) = (t, x, y, z)$ such that*

$$F = \alpha + \beta = \sqrt{|-\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2|} + b_\mu \mathrm{d}x^\mu, \tag{8.23}$$

  with b_μ = constant.

The second class of solutions in the theorem may be viewed as a Randers analog of the Bogoslovsky/*very special relativity* line element [113, 114, 115, 116]. The first class, on the other hand, may be viewed as a Randers analog of the Finsler pp-waves of m-Kropina type [43], and to a large extent also as an analog of the *very general relativity* spacetimes introduced in [44].

8.3 Modified Randers metrics

In the preceding section, we discussed solutions of Randers type, with Finsler metric $F = \alpha + \beta$. We will argue below, however, that in order to obtain a physically natural causal structure (including, for instance, both a forward and a backward light cone), the conventional definition of the Randers metric must be modified slightly. As we will see, this will not impact the characterization of exact solutions at all. The content in this section is based mostly on [H5].

After a heuristic argument that motivates the desired modification, we show that the modified Randers metric has, in particular, a very satisfactory—even pseudo-Riemannian—causal structure. As a result of this, a clear (future and

past) timelike cone can be identified and within these timelike cones the signature of the Fundamental tensor is Lorentzian everywhere. The only constraint is that $|b|^2 \equiv a^{\mu\nu}b_\mu b_\nu > -1$, which, interestingly, is in some sense the opposite[2] of the condition $|b|^2 < 1$ that appears in the well-known positive definite case, see e.g. [110]. Note that in our case, $|b|^2$ is allowed to be positive as well as negative, which is one of the properties—others will be discussed below—that distinguishes the modified Randers metrics from standard Randers metrics, for which the constraint is given by[3] $-1 < |b|^2 \leq 0$ [111].

In contrast to the preceding sections and the rest of Part III, we will momentarily set aside our assumption that $\dim M = n = 4$ and allow our manifold to have any dimension.

8.3.1 Motivation and definition

First of all, let us review why it is not that obvious in Lorentzian signature what the 'most natural' definition of the Randers metric should be, starting at the very beginning. The original definition of a Randers metric, in positive definite Finsler geometry, is that it is just a Finsler metric of the form $F = \alpha + \beta$, where $\alpha = \sqrt{a_{ij}y^i y^j}$ is a Riemannian metric and $\beta = b_i y^i$ any 1-form. This is well-defined as long as a_{ij} is positive-definite because in that case $A \equiv a_{ij}y^i y^j$ is always positive for nonvanishing vectors y. If we allow $a_{\mu\nu}$ to be a Lorentzian metric, however, the quantity A can become negative, in which case $\sqrt{A}$ is ill-defined, as we want F to be a real function. One way to remedy this, at least at a technical level, is to simply restrict the domain $\mathcal{A} \subset TM_0$ to those vectors for which $a_{\mu\nu}y^\mu y^\nu > 0$ (or alternatively, $a_{\mu\nu}y^\mu y^\nu < 0$). This is the approach that was taken in [H4] and Proposition 8.2.1 can be interpreted in that context[4]. However, restricting $\mathcal{A}$ in this way leads to issues when it comes to the physical interpretation, so here we take a different approach.

The obvious first alternative to restricting $\mathcal{A}$ to vectors with positive norm is to simply replace A by $|A|$ and define $\alpha = \sqrt{|A|}$, as we have done throughout this dissertation. Such a definition still has the side effect that F is not smooth on vectors satisfying $A = 0$, but the upside is that there's no need *a priori*

[2] If one were to adopt the opposite signature convention to ours, however, the constraint in the Lorentzian case would also turn out to be $|b|^2 < 1$, matching the positive definite case.

[3] This constraint applies to the situation where the signature is required to be Lorentzian within the whole timelike cone. It should not be confused with the situation corresponding e.g. to Proposition 8.2.1 where the region with Lorentzian signature might be strictly smaller.

[4] Since the signature convention in [H4] is opposite to the one employed here, the condition $a_{\mu\nu}y^\mu y^\nu > 0$ in that case precisely select the timelike vectors. With our current conventions, however, the relevant condition is $a_{\mu\nu}y^\mu y^\nu < 0$ as in Proposition 8.2.1.

to restrict $\mathcal{A}$ to the timelike cone of α anymore. This modification leads to a Randers metric of the form $F = \sqrt{|A|} + \beta$, which is just the standard Randers metric as we have defined it earlier. Mathematically it clearly is the most natural extension of the original definition.

From a physics perspective, however, this definition has the undesirable consequence that light rays can only propagate into one half of the tangent space, namely the half given by $\beta \leq 0$. This follows immediately from the null condition $F = 0$. Already in the simplest case—the one in which the 1-form is timelike—this leads to a situation where the light cone separates the tangent space into only two rather than three connected components[5], resulting in only one timelike cone, not both a forward *and* a backward one. Consequently, there is not a straightforward interpretation in terms of (future *and* past) timelike, spacelike, and lightlike directions, at least not in the conventional way[6].

We therefore take the viewpoint that outside of the half-space $\beta \leq 0$ in each tangent space, this version of the Randers metric is not valid, but needs to be modified. It turns out that it is possible to remove the constraint $\beta \leq 0$—extending the lightcone to the other half-space $\beta > 0$—by changing F into $F = \mathrm{sgn}(A)\sqrt{|A|} + \mathrm{sgn}(\beta)\beta = \mathrm{sgn}(A)\sqrt{|A|} + |\beta|$. The result of this is, effectively, that under some mild assumptions (details will follow below), the single light cone (from the $\beta \leq 0$ half space) is mirrored to the complementary $\beta \geq 0$ half space, whereas in the cone consisting of α-timelike vectors—the domain of the standard Randers metric (cf. Proposition 8.2.1)— F reduces to the standard Randers metric with an overall minus sign, $F = -(\alpha + \beta)$, by Proposition 8.2.1. The overall minus sign is not of any relevance and may simply be removed as a matter of convention. In particular, F is now reversible, i.e. invariant under $y \to -y$.

The preceding heuristic argument motivates the following definition.

Definition 8.3.1. *A modified Randers metric is a Finsler metric of the form,*

$$F = sgn(A)\alpha + |\beta|, \tag{8.24}$$

where we recall for completeness that $\alpha = \sqrt{|A|}$, $A = a_{\mu\nu}y^\mu y^\nu$, $\beta = b_\mu y^\nu$.

It is worth pointing out that we could have chosen a minus sign instead of a plus sign in the definition of the modified Randers metric, which would have been just as natural. But it turns out that in that case the resulting Finsler metric

[5] This can be checked easily in suitable coordinates adapted to β.

[6] We note that in some approaches a single, future pointing (by definition) cone is deemed to be sufficient [97, 19].

would not be guaranteed to have Lorentzian signature everywhere inside of the timelike cones[7]. The plus variant does have this property (as long as $|b|^2 > -1$), as we will discuss in detail below.

8.3.2 Causal structure

Our task is now to show that the modified Randers metric (8.24) indeed has very nice properties, starting with its causal structure, i.e. the structure of its light cone. By definition, the light cone is given by

$$F = 0 \qquad \Leftrightarrow \qquad \operatorname{sgn}(A) \leq 0 \ \wedge \ |A| = |\beta|^2 \qquad \Leftrightarrow \qquad A = -\beta^2, \tag{8.25}$$

and thus it follows that

$$F = 0 \qquad \Leftrightarrow \qquad (a_{\mu\nu} + b_\mu b_\nu)\mathrm{d}x^\mu \mathrm{d}x^\nu = 0, \tag{8.26}$$

meaning that the light cone of F is given precisely by the light cone of $\tilde{a}_{\mu\nu} = a_{\mu\nu} + b_\mu b_\nu$, which is itself a Lorentzian metric provided that $|b|^2 > -1$, as shown by the following Lemma.

Lemma 8.3.2. *Let $a_{\mu\nu}$ be a Lorentzian metric and b_μ a 1-form such that $|b|^2 = a_{\mu\nu}b^\mu b^\nu > -1$. Then $\tilde{a}_{\mu\nu} = a_{\mu\nu} + b_\mu b_\nu$ is also a Lorentzian metric.*

Proof. We have to show that $\tilde{a}_{\mu\nu}$ has Lorentzian signature. First, the matrix determinant lemma for rank-one updates (see e.g. [117]) says that

$$\det \tilde{a} = (1 + |b|^2) \det a. \tag{8.27}$$

Since $|b|^2 > -1$ this implies that $\det \tilde{a}$ has the same sign as $\det a$, and since $a_{\mu\nu}$ is Lorentzian, this implies $\tilde{a}$ has negative determinant everywhere. In 4D this immediately implies that $\tilde{a}$ is Lorentzian (although the signs of the eigenvalues might be flipped with respect to $a_{\mu\nu}$), but we'll allow the dimensionality to be arbitrary here. Consider the continuous family of 1-forms $b^{(\eta)}_\mu = \eta b_\mu$, where $\eta \in [0, 1]$, and define $a^{(\eta)}_{\mu\nu} = a_{\mu\nu} + b^{(\eta)}_\mu b^{(\eta)}_\nu$. For any η we have

$$\det a^{(\eta)} = \left[1 + |b^{(\eta)}|^2\right] \det a = \left[1 + \eta^2 |b|^2\right] \det a, \tag{8.28}$$

[7] In case one employs the opposite signature convention $(+,-,-,-)$ the converse would be true. In that case the preferable choice would be $F = \operatorname{sgn}(A)\alpha - |\beta|$ rather than $F = \operatorname{sgn}(A)\alpha + |\beta|$.

so $\det a^{(\eta)} \neq 0$ and hence it follows by Lemma 8.2.2 that $a^{(\eta)}$ has the same signature for all values of η. In particular, $\tilde{a} = a^{(1)}$ has the same signature as $a = a^{(0)}$, which is Lorentzian signature. □

The fact that, as long as $|b|^2 > -1$, the light cone of F coincides with the light cone of a pseudo-Riemannian metric, implies that this light cone separates the tangent space at each point into three connected components, which we can naturally interpret as the forward timelike cone, backward timelike cone, and the remainder consisting of spacelike vectors. Coincidentally we note that

$$F < 0 \qquad \Leftrightarrow \qquad (a_{\mu\nu} + b_\mu b_\nu) y^\mu y^\nu < 0, \tag{8.29}$$

from which it also follows that

$$F > 0 \qquad \Leftrightarrow \qquad (a_{\mu\nu} + b_\mu b_\nu) y^\mu y^\nu > 0. \tag{8.30}$$

This leads to the additional convenience that F-timelike vectors are precisely given by $F < 0$, and F-spacelike vectors by $F > 0$, in addition to the null vectors that are given, by definition, by $F = 0$. We summarize these results in the following proposition.

Proposition 8.3.3. *As long as $|b|^2 > -1$, the causal structure of the modified Randers metric $F = sgn(A)\alpha + |\beta|$ is identical to the causal structure of the Lorentzian metric $a_{\mu\nu} + b_\mu b_\nu$, with null vectors given by $F = 0$, timelike vectors given by $F < 0$, and spacelike vectors by $F > 0$.*

As a result of these nice features of the causal structure of the modified Randers metric, it is possible to define time orientations in a way similar to how they are defined in general relativity, namely by means of a nowhere vanishing timelike vector field T. Such T selects one of the two timelike cones as the 'forward' one, namely the one that contains T. Then another timelike vector y is future-oriented (i.e. lies in the forward timelike cone) if and only if $(a_{\mu\nu}+b_\mu b_\nu)T^\mu y^\nu < 0$. We note that similar characterizations of time orientations in Finsler spacetimes (although not in terms of an auxiliary pseudo-Riemannian metric, like $a_{\mu\nu}+b_\mu b_\nu$) have been discussed in [97].

In the special case that β is covariantly constant with respect to α (equivalently, F is Berwald)—a case that is interesting from the point of view of exact solutions—we have even stronger results. In that case, not only the causal structure but also the affine structure (i.e. affine connection) of F can be understood in terms of $a_{\mu\nu} + b_\mu b_\nu$. In order to prove this, we start by characterizing the affine structure of the latter.

Lemma 8.3.4. *Assuming that $|b|^2 > -1$, the Christoffel symbols of $\tilde{a}_{\mu\nu} = a_{\mu\nu} + b_\mu b_\nu$ can be expressed as*

$$\widetilde{\Gamma}^\rho_{\mu\nu} = \mathring{\Gamma}^\rho_{\mu\nu} + \frac{1}{1+|b|^2} b^\rho \mathring{\nabla}_{(\mu} b_{\nu)} - \left(a^{\rho\lambda} - \frac{1}{1+|b|^2} b^\rho b^\lambda \right) \left(b_\mu \mathring{\nabla}_{[\lambda} b_{\nu]} + b_\nu \mathring{\nabla}_{[\lambda} b_{\mu]} \right),$$

in terms of the Christoffel symbols $\mathring{\Gamma}^\rho_{\mu\nu}$ and Levi-Civita connection $\mathring{\nabla}$ of $a_{\mu\nu}$.

Before proving this, we point out the following immediate corollary[8].

Corollary 8.3.5. *If $\mathring{\nabla}_\mu b_\nu = 0$ then $\widetilde{\Gamma}^\rho_{\mu\nu} = \mathring{\Gamma}^\rho_{\mu\nu}$.*

Proof of Lemma 8.3.4. It can easily be checked that the inverse of $\tilde{a}_{\mu\nu}$ is given by

$$\tilde{a}^{\mu\nu} = a^{\mu\nu} - \frac{1}{1+|b|^2} b^\mu b^\nu. \tag{8.31}$$

Unless otherwise specified (as in the case of $\widetilde{\Gamma}$ below!) indices are raises and lowered with $a_{\mu\nu}$. Writing $\mathring{\Gamma}_{\lambda\mu\nu} = a_{\lambda\rho}\mathring{\Gamma}^\rho_{\mu\nu}$ and $\widetilde{\Gamma}_{\lambda\mu\nu} = \tilde{a}_{\lambda\rho}\widetilde{\Gamma}^\rho_{\mu\nu}$ we first note that we can express the latter as

$$\begin{aligned}
\widetilde{\Gamma}_{\lambda\mu\nu} &= \tfrac{1}{2}\left(\partial_\mu \tilde{a}_{\lambda\nu} + \partial_\nu \tilde{a}_{\mu\lambda} - \partial_\lambda \tilde{a}_{\mu\nu}\right) && (8.32)\\
&= \mathring{\Gamma}_{\lambda\mu\nu} + b_\lambda \partial_{(\mu} b_{\nu)} - b_\mu \partial_{[\lambda} b_{\nu]} - b_\nu \partial_{[\lambda} b_{\mu]}, && (8.33)
\end{aligned}$$

where (μ,ν) denotes symmetrization and $[\mu,\nu]$ denotes anti-symmetrization. It thus follows that

$$\begin{aligned}
\widetilde{\Gamma}^\rho_{\mu\nu} &= \tilde{a}^{\rho\lambda}\widetilde{\Gamma}_{\lambda\mu\nu} && (8.34)\\
&= \left(a^{\rho\lambda} - \frac{1}{1+|b|^2} b^\rho b^\lambda \right) \left(\mathring{\Gamma}_{\lambda\mu\nu} + b_\lambda \partial_{(\mu} b_{\nu)} - b_\mu \partial_{[\lambda} b_{\nu]} - b_\nu \partial_{[\lambda} b_{\mu]} \right) && (8.35)\\
&= \mathring{\Gamma}^\rho_{\mu\nu} - \frac{1}{1+|b|^2} b^\rho b_\lambda \mathring{\Gamma}^\lambda_{\mu\nu} + \left(a^{\rho\lambda} - \frac{1}{1+|b|^2} b^\rho b^\lambda \right) b_\lambda \partial_{(\mu} b_{\nu)} && (8.36)\\
&\qquad - \left(a^{\rho\lambda} - \frac{1}{1+|b|^2} b^\rho b^\lambda \right) \left(b_\mu \partial_{[\lambda} b_{\nu]} + b_\nu \partial_{[\lambda} b_{\mu]} \right). && (8.37)
\end{aligned}$$

[8] This corollary can also be proven without reference to Lemma 8.3.4, by simply noticing that $\mathring{\nabla}_\rho \tilde{a}_{\mu\nu} = 0$, which implies that $\mathring{\nabla}$ must also be the Levi-Civita connection of $\tilde{a}_{\mu\nu}$.

The second and third term add up to

$$-\frac{1}{1+|b|^2}b^\rho b_\lambda \mathring{\Gamma}^\lambda_{\mu\nu} + \left(a^{\rho\lambda} - \frac{1}{1+|b|^2}b^\rho b^\lambda\right) b_\lambda \partial_{(\mu} b_{\nu)} \tag{8.38}$$

$$= -\frac{1}{1+|b|^2}b^\rho b_\lambda \mathring{\Gamma}^\lambda_{\mu\nu} + b^\rho \partial_{(\mu} b_{\nu)} - \frac{|b|^2}{1+|b|^2}b^\rho \partial_{(\mu} b_{\nu)} \tag{8.39}$$

$$= -\frac{1}{1+|b|^2}b^\rho b_\lambda \mathring{\Gamma}^\lambda_{\mu\nu} \; + \frac{1}{1+|b|^2}b^\rho \partial_{(\mu} b_{\nu)} \tag{8.40}$$

$$= \frac{1}{1+|b|^2}b^\rho \left(\partial_{(\mu} b_{\nu)} - b_\lambda \mathring{\Gamma}^\lambda_{\mu\nu}\right) \tag{8.41}$$

$$= \frac{1}{1+|b|^2}b^\rho \nabla_{(\mu} b_{\nu)}, \tag{8.42}$$

which results in

$$\widetilde{\Gamma}^\rho_{\mu\nu} = \mathring{\Gamma}^\rho_{\mu\nu} + \frac{1}{1+|b|^2}b^\rho \nabla_{(\mu} b_{\nu)} - \left(a^{\rho\lambda} - \frac{1}{1+|b|^2}b^\rho b^\lambda\right)\left(b_\mu \partial_{[\lambda} b_{\nu]} + b_\nu \partial_{[\lambda} b_{\mu]}\right). \tag{8.43}$$

Finally, we may replace all partial derivatives with covariant ones because

$$2\nabla_{[\lambda} b_{\nu]} = \nabla_\lambda b_\nu - \nabla_\nu b_\lambda = \partial_\lambda b_\nu - \partial_\nu b_\lambda = 2\partial_{[\lambda} b_{\nu]}. \tag{8.44}$$

That yields the desired formula. □

Theorem 8.3.6. *If β is covariantly constant with respect to α and satisfies $|b|^2 > -1$ then the causal structure and the affine structure of the modified Randers metric $F = sgn(A)\alpha + |\beta|$ are identical to those of the Lorentzian metric $\tilde{a}_{\mu\nu} = a_{\mu\nu} + b_\mu b_\nu$. In other words, the timelike, spacelike, and null geodesics of F coincide with the timelike, spacelike, and null geodesics, respectively, of $\tilde{a}_{\mu\nu}$.*

Proof. Proposition 8.3.3 guarantees that the causal structures coincide. By Proposition 5.3.4 it follows that the affine structure of F coincides with that of α and Corollary 8.3.5 shows that this affine structure is the same as that of $\tilde{a}_{\mu\nu}$. □

We note that in the proof of Theorem 8.3.6, we have applied Proposition 5.3.4, which is a result for standard Randers metrics. Strictly speaking, however, the modified Randers metric F is not even an (α, β)-metric, let alone a Randers metric. This is because F depends not only on α and β but also on the sign

of A. The following argument shows that Proposition 5.3.4 is nevertheless valid also in this case.

If we denote by $\mathcal{A}$ the maximal conic subbundle of TM on which F satisfies all axioms of a Finsler metric then we have $\mathcal{A} \subset \mathcal{A}_{++} \cup \mathcal{A}_{+-} \cup \mathcal{A}_{-+} \cup \mathcal{A}_{--}$ in terms of the conic subbundles

$$\mathcal{A}_{\pm\pm} = \{(x, y) \in \mathcal{A} : \pm A > 0, \pm\beta > 0\}, \tag{8.45}$$

where the first (second) $\pm$ on the LHS refers to the first (second) $\pm$ on the RHS. This decomposition holds because F is differentiable neither at $A = 0$ nor at $\beta = 0$, so that these zero sets are not contained in $\mathcal{A}$. On each of the subdomains $\mathcal{A}_{\pm\pm}$, F is just a standard Randers metric (perhaps up to an irrelevant overall minus sign) and this is why many of the results for standard Randers metrics carry over to modified Randers metrics. This is true in particular for Proposition 5.3.4 and hence the proof of Theorem 8.3.6 is perfectly valid. We will come back to this in the context of exact solutions later.

Apart from the intrinsic significance of Theorem 8.3.6, one of its nice consequences is the existence of so-called radar neighborhoods [118]. In words, this means that, given any observer and any spacetime point sufficiently close to the observer's worldline, there is (at least locally) exactly one future-pointing light ray and one past-pointing light ray that connect that point to the worldline of the observer. The precise statement is as follows.

Proposition 8.3.7 (Existence of radar neighborhoods)**.** *Consider a Finsler spacetime with modified Randers metric of Berwald type and with* $|b|^2 > -1$*, let* γ *be a timelike curve and let* $p = \gamma(\lambda_0)$ *be some point on* γ*. Then there are open subsets* U *and* V *with* $p \in U \subset V$ *such that every point* $q \in U \setminus \mathrm{image}(\gamma)$ *can be connected to* γ *by precisely one future-pointing and precisely one past-pointing null geodesic that stays within* V*.*

Proof. This result has been proven in [118] for spacetimes with a Lorentzian metric rather than a modified Randers metric. However, the statement relies only on the properties of null geodesics and timelike curves. By Theorem 8.3.6, null geodesics and timelike curves of the modified Randers metric can be understood as null geodesics and timelike curves, respectively, of some Lorentzian metric (note that F being Berwald implies that $\mathring{\nabla}_\mu b_\nu = 0$, by Proposition 5.3.4). Hence the result extends immediately to our case. □

The set U is called a radar neighborhood of p with respect to γ. The existence of radar neighborhoods will play an important role in Chapter 9 when we calculate

the radar distance to a given spacetime point during the passing of a Finslerian gravitational wave of modified Randers type and compare it to the result for a classical general relativistic gravitational wave. Proposition 8.3.7 essentially says that this radar distance is a well-defined notion to begin with.

8.3.3 Signature and regularity

Our next objective is to determine the signature and regularity (i.e. degree of smoothness) of F. Provided α is Lorentzian, we will find the signature of F to be Lorentzian everywhere within the timelike cone, except (trivially) at those points where F is degenerate or not smooth, i.e. at the irregularities of F. This holds irrespective of the properties of the 1-form. Therefore we will be interested in the subset of the timelike cone where F is not regular. Ideally, this subset is empty, and we will see that when β is timelike or null this is indeed the case.

Signature of the fundamental tensor

The determinant of the fundamental tensor of an (α, β)-metric $F = \alpha\phi(s)$, with $s = \beta/\alpha$ is given by

$$\det g = \phi^{n+1}(\phi - s\phi')^{n-2}(\phi - s\phi' + (\operatorname{sgn}(A)|b|^2 - s^2)\phi'') \det a, \tag{8.46}$$

where $n = \dim M$ and the prime denotes differentiation with respect to s. This formula, which appeared originally in [H5], is a generalization of a well-known result for positive definite (α, β)-metrics, and its proof can be found in Appendix C. Because of the appearance of $\operatorname{sgn}(A)$ the expression is slightly different from its positive definite counterpart, to which it reduces when $A > 0$, i.e. $\operatorname{sgn}(A) = 1$. For a modified Randers metric of the form $F = \operatorname{sgn}(A)\alpha + |\beta|$ the function ϕ is given by $\phi(s) = \operatorname{sgn}(A) + |s|$, so (8.46) reduces to

$$\frac{\det g}{\det a} = \operatorname{sgn}(A)^{n-1}\,(\operatorname{sgn}(A) + |s|)^{n+1} = \left(\operatorname{sgn}(A)\frac{F}{\alpha}\right)^{n+1}. \tag{8.47}$$

We denote by $\mathcal{A}$ the maximal conic subbundle of TM on which F satisfies all axioms of a Finsler metric. The precise characteristics of $\mathcal{A}$ are yet to be determined, but since F is differentiable neither at $\alpha = 0$ nor at $\beta = 0$, $\mathcal{A}$ can clearly only contain points $(x, y) \in TM_0$ that satisfy $\alpha \neq 0$ and $\beta \neq 0$. We will tacitly use this fact in what follows.

Lemma 8.3.8. *In even dimension n, the subset of $\mathcal{A}$ on which g has Lorentzian signature is given precisely by those points in $\mathcal{A}$ that satisfy $\operatorname{sgn}(A)F > 0$.*

Proof. Consider the family of functions $F_\lambda = (1-\lambda)\mathrm{sgn}(A)\alpha + \lambda F = \mathrm{sgn}(A)\alpha + \lambda|\beta|$ for $\lambda \in [0,1]$ and suppose first that $\mathrm{sgn}(A)F > 0$. Then both $\mathrm{sgn}(A)F_0 = \alpha > 0$ and $\mathrm{sgn}(A)F_1 = \mathrm{sgn}(A)F > 0$, and since $F_\lambda = (1-\lambda)F_0 + \lambda F_1$ is a convex sum, this implies that $\mathrm{sgn}(A)F_\lambda > 0$ for each λ. Since we also know that $\det a < 0$, (8.47) then implies that $\det g_\lambda$ has the same sign for each λ and hence it follows from Lemma 8.2.2 that the signature of $g = g_1$ is the same as that of $\epsilon a = g_0$, i.e. Lorentzian.
Conversely, if $\mathrm{sgn}(A)F \leq 0$ then (8.47) shows that $\det g \geq 0$, as $n+1$ is assumed to be odd. Hence in this case g is certainly not Lorentzian. □

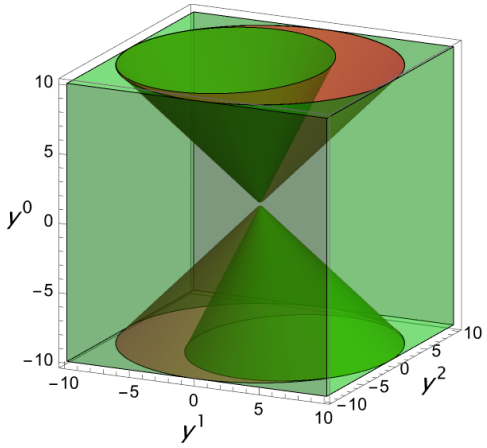
Let us investigate what this set $\mathrm{sgn}(A)F > 0$ looks like. First note that $F < 0$ trivially implies $A < 0$, meaning that the F-timelike cone (recall Proposition 8.3.3) lies completely within the A-timelike cone. It follows that if $F < 0$ then $\mathrm{sgn}(A)F > 0$ and hence the region $F < 0$ has Lorentzian signature. In other words, within the entire timelike cone of F (intersected with $\mathcal{A}$), the signature of the fundamental tensor is Lorentzian. Similarly, $A > 0$ implies $F > 0$ and hence $\mathrm{sgn}(A)F > 0$, so the region $A > 0$ also has Lorentzian signature. What remains is the region where simultaneously $A \leq 0$ and $F \geq 0$, and in this case, it is clear that $\mathrm{sgn}(A)F \leq 0$. Hence, apart from possible irregularities, this is precisely the subset of TM where the signature is *not* Lorentzian. The inequalities describing this region can be interpreted as saying that y lies within the A-timelike cone but outside of the F-timelike cone (or possibly on the boundary of either). In other words, this is precisely the region in between the light cone of A and that of F. We thus conclude that F has Lorentzian signature everywhere on $\mathcal{A}$ except in the region in between the two light cones (including their boundaries[9]).

Figure 8.1 shows a $(1+2)$D projection of the light cone of the modified Randers metric, and it also shows the region (in green) where the signature of the fundamental tensor is Lorentzian, for each possible causal character of the 1-form. In each subfigure, the inner light cone is that of F and the outer light cone that of A.

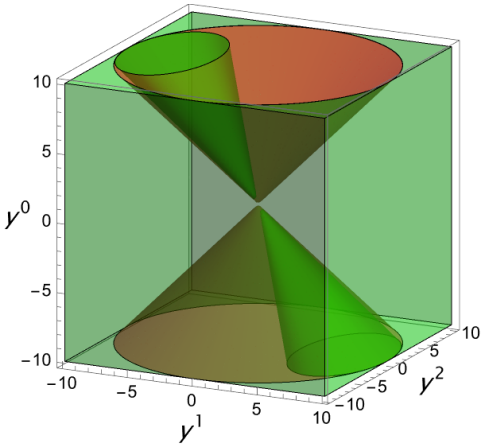
Regularity

From the definition of the modified Randers metric, it is clear that the set of points where F (or more relevantly, F^2) is *not* smooth is given by $A = 0$ or $\beta = 0$. We would like to know which of these points, if any, lie within the timelike

[9] On the light cone of A the Finsler metric is not even differentiable, so it certainly does not have Lorentzian signature there. On the light cone of F, either $A = 0$, in which case the previous comment applies, or $A \neq 0$, in which case $\det g = 0$, by (8.47), so that the point does not lie in $\mathcal{A}$.

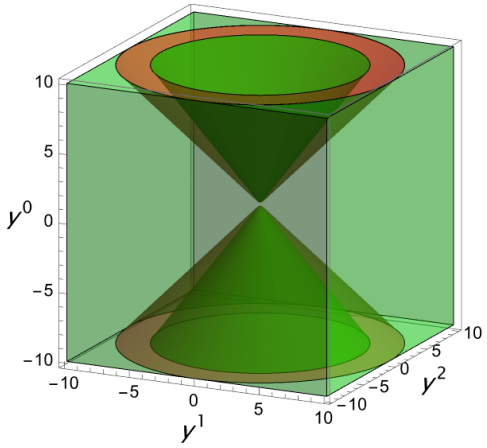


(8.1a) Null 1-form, $\rho = 0.6$

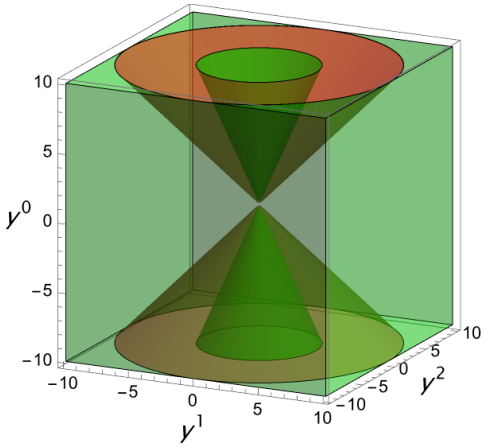


(8.1b) Null 1-form, $\rho = 1.4$

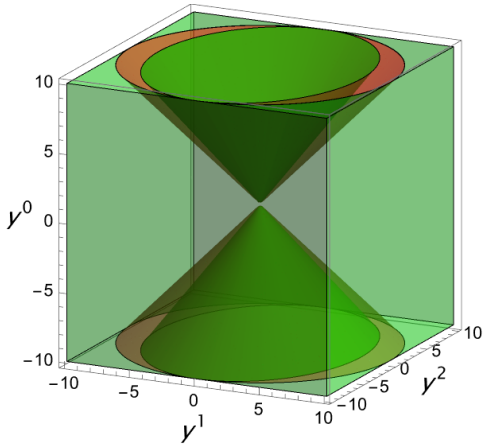
Figure 8.1: The figures show, for several representative values of ρ, a 3D projection of the light cone and the signature of the fundamental tensor of the modified Randers metric $F = \mathrm{sgn}(A)\alpha + |\beta|$, in coordinates such that at $x \in M$ one has $A = -(y^0)^2 + (y^1)^2 + \cdots + (y^2)^{n-1}$ and $\beta = \rho\, y^0$ (in the timelike case) or $\beta = \rho(y^0 + y^1)$ (in the null case) or $\beta = \rho\, y^1$ (in the spacelike case). Such coordinates can always be chosen. Green regions correspond to Lorentzian signature and red regions to non-Lorentzian signature. Figures 8.1a - 8.1f show the physically reasonable scenarios, where $|b|^2 > -1$. In that case, two cones can be observed. The inner cone is the true light cone of F (i.e. the set $F = 0$), and the outer cone is the light cone of a_{ij} (i.e. the set $A = 0$). The only region with non-Lorentzian signature (disregarding irregularities) is precisely the gap in between the two cones, including their boundaries. If on the other hand $|b|^2 = -1$ (Figure 8.1g) then the light 'cone' of F is the line $y^1 = y^2 = y^3 = 0$. And if $|b|^2 < -1$ (Figure 8.1h) then the light 'cone' of F consists only of the origin. Therefore we deem the latter two cases not physically interesting.



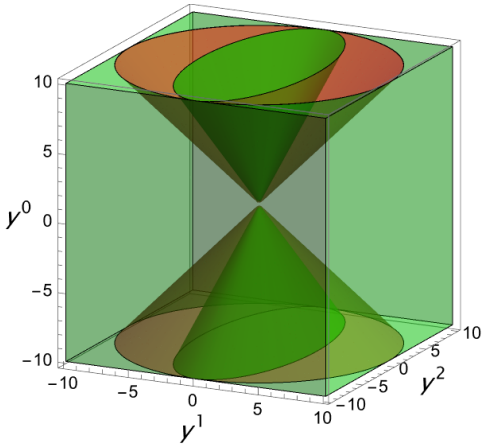
(8.1c) Timelike 1-form, $\rho = 0.65$



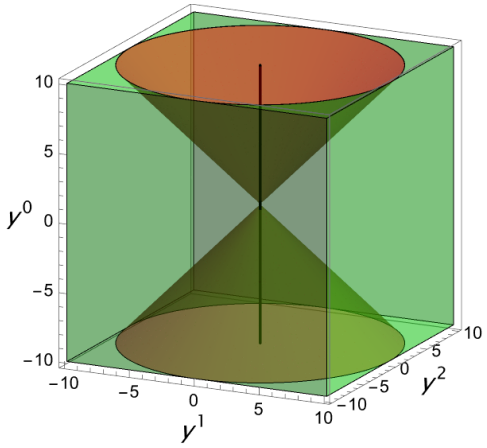
(8.1d) Timelike 1-form, $\rho = 0.9$



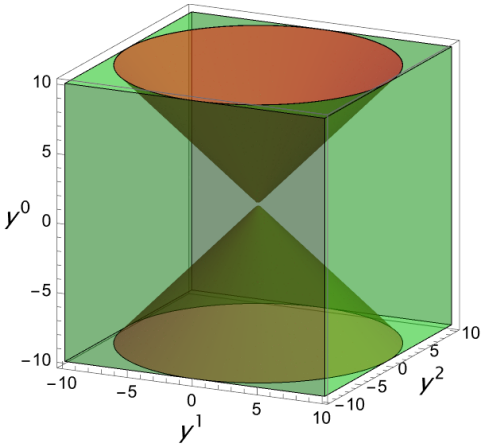
(8.1e) Spacelike 1-form, $\rho = 0.8$



(8.1f) Spacelike 1-form, $\rho = 2$



(8.1g) Timelike 1-form, $|b|^2 = -1$.



(8.1h) Timelike 1-form, $|b|^2 < -1$.

cone of F, or perhaps on its boundary, the light cone. First, we focus on the possibility that $A = 0$. As observed above and confirmed visually in Figure 8.1, the F-timelike cone lies completely within the A-timelike cone, which implies that none of the $A = 0$ irregularities lie within the F-timelike cone. But they might lie on the light cone, namely if the light cone of A and the light cone of F have a nonempty intersection. The specifics of this are straightforward to derive, but depend on the (pointwise) causal character of the 1-form β; we list the results below and leave the details to the reader. The easiest way to check these properties is to use coordinates in which

$$A = -(\mathrm{d}x^0)^2 + \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j, \quad \beta = \begin{cases} \rho\,\mathrm{d}x^0 & \textit{timelike} \\ \rho\,(\mathrm{d}x^0 + \mathrm{d}x^1) & \textit{null} \\ \rho\,\mathrm{d}x^1 & \textit{spacelike} \end{cases} \tag{8.48}$$

at a given point $x \in M$. Such coordinates always exist[10]. In each of the following cases, we consider the light cone(s) at the fixed point x.

[10]We recall that this can be seen as follows. First, since a_{ij} is Lorentzian, it is always possible to choose coordinates such that A is just the Minkowski metric at a given point $x \in M$. Writing $b^\mu = (b^0, b^1, \dots b^{n-1})$ in these coordinates, we may apply a spatial rotation to the coordinates $b^1, \dots, b^{n-1}$, such that they are transformed into $(b^1, 0, \dots, 0)$, leaving the metric at x unchanged. Then $b^\mu = (b^0, b^1, 0, \dots, 0)$. Now we separate the three cases. If $|b|^2 = 0$, it follows that $b^1 = \pm b^0$ and by applying—if necessary—a spatial reflection in the x^1 direction we may choose either sign. If $|b|^2 < 0$ then we may go to the local rest frame by a Lorentz transformation, resulting in $b^1 = 0$. If on the other hand $|b|^2 > 0$ we may perform a Lorentz transformation resulting in $b^0 = 0$.

- If β is timelike at x and $|b|^2 > -1$ then the light cones do not intersect (apart from the trivial intersection in the origin). Hence not only the F-timelike cone but also its boundary, the light cone, lies completely within the (interior of the) light cone of A.
- If β is null at x then the intersection of the two lightcones is given by those y^μ that are multiples of b^μ. This intersection thus spans a single line in the tangent space.
- If β is spacelike at x (and assuming $\dim M > 2$) then a_{ij} induces an $(n-1)$-dimensional Lorentzian metric on the hypersurface $\beta = 0$. The two light cones intersect along the light cone of this induced Lorentzian metric. In Figure 8.1e and Figure 8.1f, which are 3D projections, the latter is given by $y^0 = \pm y^2$; hence the intersection appears as a pair of lines.
- If $|b|^2 = -1$ then the light cone of F is given by a single line, namely the line consisting of all multiples of b^μ. In Figure 8.1g this is just the y^0-axis.
- If $|b|^2 < -1$ then the light cone of F is given by the single point $y = 0$.

This sums up the $A = 0$ irregularities on the light cone of F. Regarding the $\beta = 0$ irregularities, note that $\beta = a_{\mu\nu} b^\mu y^\mu = 0$ if and only if y^μ is orthogonal to b^μ. The orthogonal complement to a timelike vector is spacelike, so if b_μ is timelike, β cannot vanish on A-timelike vectors and hence in particular $\beta \neq 0$ in the F-timelike cone. If b_μ is null, on the other hand, then the only nonspacelike vectors orthogonal to it are multiples of b^μ itself. These coincide with the $A = 0$ irregularities in the second bullet point above. Only in the case that b_μ is spacelike will there be $\beta = 0$ irregularities inside the F-timelike cone, namely all A-timelike vectors orthogonal to b_μ.

Now let us combine everything we have worked out in this section so far. A modified Randers metric with $|b|^2 > -1$ has Lorentzian signature everywhere within its timelike cone, with the exception of possible irregularities in the cone. The latter are given either by $\beta = 0$, characterized in the preceding paragraph, or by $A = 0$, corresponding to one of the intersections listed in the bullet list above, depending on the causal character of β. We may thus summarize the results as follows, focusing for clarity on the case where the 1-form has a global causal character, i.e. is everywhere timelike, spacelike or null.

Regularity and Signature - Summary

Let F be a modified Randers metric with $|b|^2 > -1$.

- If β is **timelike** at then F is smooth and has Lorentzian signature within its timelike cone. On the light cone, F is smooth but degenerate.
- If β is **null** then F is smooth and has Lorentzian signature within its timelike cone. On the light cone, F is degenerate and there is a 1–dimensional irregularity where F is not smooth, namely the line spanned by b^μ.
- If β is **spacelike** then within its timelike cone, F has Lorentzian signature but has an $(n-1)$-dimensional irregularity where F is not smooth, given by the hyperplane $\beta = 0$. On the light cone, F is degenerate and there is an $(n-2)$-dimensional irregularity where F is not smooth, given by the light cone of the induced metric on the submanifold $\beta = 0$.

If $|b|^2 \leq -1$ the timelike cone is the empty set, suggesting that this case is not physically relevant.

Clearly, the timelike case is the winner in terms of desirable properties from the point of view of physics. But the null case is a close runner-up, and even the spacelike scenario is acceptable. In particular, the set where F is not smooth has measure zero and empty interior in each of the cases, provided $\dim M > 2$. This is true even when the 1-form is spacelike. On the other hand, the *classical* Randers metric can only be considered a physically reasonable Finsler spacetime if the 1-form is either null or timelike [111], and even in these cases, there is only a future timelike cone (or only a past timelike cone, but not both). This motivates the use of modified Randers metrics rather than standard Randers metrics in Finsler gravity.

8.3.4 Exact vacuum solutions

As noted already below Theorem 8.3.6, a modified Randers metric F is 'locally' just a standard Randers metric, in the sense that the domain $\mathcal{A}$ of F can be decomposed into several parts, on each of which F reduces to a standard Randers metric. That means that on each of the individual subdomains, all results obtained in Section 8.2 are applicable. And since the classification results pertaining to exact solutions do not depend on the specifics of the domain, it follows that the results remain true also for modified Randers metrics considered on their entire domain $\mathcal{A}$. We summarize the most important ones below.

- For modified Randers metrics of Berwald type, Pfeifer and Wohlfarth's field equation is equivalent to Rutz's equation, which comes down to Ricci-flatness, $\mathrm{Ric} = 0$, or equivalently, $R_{\mu\nu} = 0$. This is the analog of Proposition 8.2.4.
- From this, it follows that a modified Randers metric $F = \mathrm{sgn}(A)\alpha + |\beta|$ of Berwald type is a vacuum solution (to any and each of these equations) if and only if α is a vacuum solution to Einstein's equations. This is the analog of Theorem 8.2.6.
- The analog of the final classification result, Corollary 8.2.8, holds for modified Randers metrics as well.

CHAPTER 9

Finsler Gravitational Waves

The aim of this chapter, which is based mostly on [H5], is to attach a physical interpretation to, and investigate the physical consequences of, the (plane) pp-wave solutions of (α, β)-metric type obtained in Section 8.1. We will achieve this by means of a two-fold linearization procedure, after which the whole class of solutions essentially reduces to a single family of Randers gravitational waves parameterized by a $+$ and a $\times$ polarization, for which it is possible to calculate exactly what happens when such a wave passes, say, a gravitational wave interferometer on Earth.

The relevant observable measured in interferometer experiments is essentially the radar distance, so we compute this radar distance for our Finslerian gravitational waves, reproducing in the appropriate limit the radar distance formula for a standard gravitational wave in general relativity (GR) [45]. Although at first sight, the expression for the Finsler radar length looks quite different from its GR counterpart, we show that this is nothing but a coordinate artifact. Remarkably, when the Finslerian expression is interpreted correctly in terms of observable quantities, it becomes clear that the two expressions are in fact identical, suggesting that there is no observational difference between the Finsler and GR case, at least as far as radar distance measurements are concerned.

This is the first example of a situation in which an explicit expression for the Finslerian radar distance could be obtained in the case of finite spacetime separations[1], and as such our results may be viewed in part as a proof of concept.

[1] In the case of infinitesimal spacetime separations the radar length has been studied in

9.1 Linearized pp-wave solutions

There are two steps to the linearization. First, we linearize in the departure from flatness by writing the Lorentzian part α (or rather A) as the flat Minkowski metric plus a small deviation, resulting in a standard general relativistic linearized gravitational wave. Second, we linearize in the departure from GR, reducing the class of (α, β)-metrics to metrics of Randers type. Our starting point will be the expression (8.10) for the plane wave solutions in Rosen coordinates.

9.1.1 Linearized gravitational wave solutions

As anticipated, we start by linearizing the plane wave metric

$$A = -2\mathrm{d}u\mathrm{d}v + h_{ij}(u)\mathrm{d}x^i\mathrm{d}x^j \tag{9.1}$$

appearing in (8.10) in its departure from flatness. In other words, we consider the scenario that α is very close to the flat Minkowski metric. In this case we may write $h_{ij}(u) = \delta_{ij} + \varepsilon f_{ij}(u)$ with $\varepsilon \ll 1$. The linearized field equations (i.e. to first order in ε) for α then read[2]

$$f''_{11}(u) + f''_{22}(u) = 0. \tag{9.2}$$

Hence f_{11} and $-f_{22}$ must be equal up to an affine function of u. Here we will focus on the case where $f_{11} = -f_{22}$, which can always be achieved by means of the transverse traceless gauge[3]. Conventionally one writes the subscripts as $f_{11} = -f_{22} =: f_+$ and $f_{12} = f_{21} =: f_\times$, denoting the plus– and cross-polarization of the gravitational wave, so we will stick to that notation from here onwards. That brings us to the Finsler metric

$$F = \alpha\,\phi(\beta/\alpha), \tag{9.3}$$
$$\begin{cases} A = -2\mathrm{d}u\mathrm{d}v + (1+\varepsilon f_+(u))\mathrm{d}x^2 + (1-\varepsilon f_+(u))\mathrm{d}y^2 + 2\varepsilon f_\times(u)\mathrm{d}x\,\mathrm{d}y \\ \beta = \mathrm{d}u \end{cases}$$

[119, 120].

[2] The (linearized) vacuum field equation for F is more complicated in general, but as discussed extensively in Section 8.1, if the vacuum field equation for α is satisfied then so is the field equation for F. In the case of Randers metrics, to which everything will reduce momentarily, the field equation for F is even equivalent to the field equations for α. Hence for our present purposes, the field equations for α suffice.

[3] We leave open the question of whether the form of the 1-form $\beta = \mathrm{d}u$ always remains invariant under such a transformation to the transverse traceless gauge.

Note that if we substitute $u = (t-z)/\sqrt{2}$ and $v = (t+z)/\sqrt{2}$, then A reduces to the standard expression for a gravitational wave metric in GR. In these coordinates, the Finsler metric reads

$$F = \alpha\,\phi(\beta/\alpha), \tag{9.4}$$
$$\begin{cases} A = -\mathrm{d}t^2 + \mathrm{d}z^2 + (1+\varepsilon f_+(t-z))\mathrm{d}x^2 + (1-\varepsilon f_+(t-z))\mathrm{d}y^2 \\ \qquad\qquad +2\varepsilon f_\times(t-z)\mathrm{d}x\,\mathrm{d}y \\ \beta = \frac{1}{\sqrt{2}}\,(\mathrm{d}t - \mathrm{d}z)\,. \end{cases}$$

We will thus refer to such Finsler metrics as (Finslerian) *gravitational wave* metrics.

9.1.2 Linearized (α,β)-metrics are Randers metrics

It is natural to linearize not only in ε, characterizing the departure from flatness, but to also use a perturbative expansion in the 'size' of the 1-form, characterizing the departure from GR and pseudo-Riemannian geometry. In other words, we consider gravitational waves that deviate only slightly from the well-known general relativistic ones. The purpose of this section is to highlight that any (α,β)-metric is equivalent to a Randers metric, to first order in this expansion, so that from a physics point of view, Randers metrics are quite a bit more general than they might seem at first glance.

So consider an (α,β)-metric constructed from a pseudo-Riemannian metric α and a 1-form β such that $\beta \ll 1$. To see what happens in such a scenario, we replace β with $\lambda\beta$ and expand to first order in λ. Then we obtain

$$F = \alpha\phi\left(\frac{\lambda\beta}{\alpha}\right) \approx \alpha\left(\phi(0) + \lambda\phi'(0)\frac{\beta}{\alpha}\right) = \alpha\phi(0) + \lambda\phi'(0)\beta = \tilde{\alpha} + \tilde{\beta}. \tag{9.5}$$

Hence to first order in λ, any (α,β)-metric is indeed equivalent to a Randers metric[4]. Consequently, by replacing $\mathrm{d}u$ by $\lambda\mathrm{d}u$ in (9.3), which technically can be achieved by a coordinate transformation that scales u by λ and v by $1/\lambda$, it follows that to first order in λ the Finsler metric of the (α,β)-type gravitational

[4] Actually, this is not true for *all* (α,β)-metrics but only those which allow an expansion around $s = \beta/\alpha = 0$. This excludes some of the m-Kropina metrics, for instance, because these are not always well-behaved in the limit $\beta \to 0$.

waves takes the form,

$$F = \alpha + \beta, \tag{9.6}$$
$$\begin{cases} A = -2\mathrm{d}u\mathrm{d}v + (1+\varepsilon f_+(u))\mathrm{d}x^2 + (1-\varepsilon f_+(u))\mathrm{d}y^2 + 2\varepsilon f_\times(u)\mathrm{d}x\,\mathrm{d}y \\ \beta = \lambda \mathrm{d}u \end{cases}$$

The parameter λ then characterizes the departure from GR and pseudo-Riemannian geometry. We will assume without loss of generality that $\lambda > 0$.

One more step is necessary to arrive at the Finsler metric that we will take as the starting point for the calculation of the radar distance in Section 9.2. And this step has to do with the causal structure of the Randers metric. We carefully argued in Section 8.3 that the physically relevant version of the Randers metric is not the standard one, but that one should prefer the *modified Randers metric*. Accordingly, the physically relevant version of the Finslerian gravitational wave metric is given by

$$F = \mathrm{sgn}(A)\alpha + |\beta|, \tag{9.7}$$
$$\begin{cases} A = -2\mathrm{d}u\mathrm{d}v + (1+\varepsilon f_+(u))\mathrm{d}x^2 + (1-\varepsilon f_+(u))\mathrm{d}y^2 + 2\varepsilon f_\times(u)\mathrm{d}x\,\mathrm{d}y \\ \beta = \lambda \mathrm{d}u \end{cases}$$

For completeness, we also include the form of the Finsler metric in the coordinates (t, x, y, z) defined right below (9.3):

$$F = \mathrm{sgn}(A)\alpha + |\beta|, \tag{9.8}$$
$$\begin{cases} A = -\mathrm{d}t^2 + \mathrm{d}z^2 + (1+\varepsilon f_+(t-z))\mathrm{d}x^2 + (1-\varepsilon f_+(t-z))\mathrm{d}y^2 \\ \qquad\qquad +2\varepsilon f_\times(t-z)\mathrm{d}x\,\mathrm{d}y \\ \beta = \frac{\lambda}{\sqrt{2}}\left(\mathrm{d}t - \mathrm{d}z\right), \end{cases}$$

Let us briefly explain why this modification of the Randers metric is perfectly allowed, even though it might seem at first glance to be in conflict with the result obtained above that any (α, β)-metric should reduce to a *conventional* Randers metric. What is important to note is that there is in principle the possibility that different regions of the tangent bundle (subdomains of $\mathcal{A}$) could house different (α, β)-metrics, hence leading to different Randers metrics in each subdomain.

More precisely, we could define one (α, β)-metric F_1 on a conic subbundle $\mathcal{A}_1 \subset TM_0$ and another (α, β)-metric, F_2, on a different conic subbundle $\mathcal{A}_2 \subset TM_0$. If the two subbundles do not overlap then this defines a perfectly valid (α, β)-type Finsler spacetime on the union $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$. Admittedly, it would not satisfy the definition of an (α, β)-metric in the strict sense of Definition 5.2.1

on its domain as a whole, but there is no reason at this point in the analysis why we should be so strict. For our present purposes, having a physically viable causal structure is clearly more important than adhering to a mathematically convenient—but arbitrary from the physics perspective—definition.

To first order in the deviation from pseudo-Riemannian geometry, the so-obtained Finsler metric would reduce to a certain Randers metric on $\mathcal{A}_1$ and to a different Randers metric on $\mathcal{A}_2$. Hence on $\mathcal{A}$ as a whole, the resulting linearized metric might not be expressible as a single standard Randers metric. This is indeed precisely the case for the modified Randers metric. On each of the four subsets of $\mathcal{A}$ given by $\pm A > 0, \pm\beta > 0$, the modified Randers metric $F = \mathrm{sgn}(A)\alpha + |\beta|$ reduces to a standard Randers metric (up to a possible overall minus sign, which is irrelevant). Our use of the modified Randers metric is therefore completely consistent with the previous results, and it is preferred here over the standard Randers metric because of its physically natural causal structure and its other desirable properties.

9.2 Observational signature

Now we are finally in the position to analyze the physical effects of a passing Finslerian gravitation wave of (α, β)-type. Our starting point will be the linearized Finsler metric (9.8). Since actual gravitational wave measurements are very often done by interferometers, which effectively measure the *radar distance*, the aim of this section is to compute this radar distance during the passing of a gravitational wave of the form (9.8).

The setup is as follows. A light ray is emitted from some spacetime location with coordinates (t_0, x_0, y_0, z_0), travels to another location in spacetime with coordinates $(t_0 + \Delta t, x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$, where it is reflected and after which it travels back to the original (spatial) location, with spacetime coordinates $(t_0+\Delta t_{\mathrm{tot}}, x_0, y_0, z_0)$, being received there again. We are interested in the amount of proper time that passes between emission and reception of the light ray, as measured by an 'inertial' observer[5] located at the spatial coordinates (x_0, y_0, z_0). Because light travels forward *and* backward during this time interval, one half of the time interval is usually called the *radar distance* between the two spacetime

[5]In this context, we say that an observer is intertial if it would be considered an intertial observer in the absence of the wave (i.e. for $f_+ = f_\times = 0$). In other words, thinking of the gravitational wave as having a finite duration as it passes the Earth, an observer is inertial precisely if it is inertial before and after the wave passes. In the absence of the wave, an observer is said to be inertial if it has constant spatial position (x, y, z) in some coordinate system in which $A = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$ and $\beta = \mathrm{d}u$.

points (in GR the value is sometimes multiplied by the velocity of light, c, which we have set to 1, so as to obtain a quantity with the dimensions of distance). In other words, the radar distance can be expressed as $R = \Delta\tau/2$. Mathematically, to obtain the radar distance (as measured by the interferometer) to the point $(t_0 + \Delta t, x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$, we have to find the unique (cf. Proposition 8.3.7) two null geodesics that connect this point to the worldline of the observer.

The expression for the radar distance of a standard gravitational wave in GR has been obtained in [45] and our calculation below follows essentially the same methods. At each step of the calculation, we will clearly point out the differences with the corresponding situation in GR, so that it is clear where each of the Finslerian effects (three separate effects can be identified) enters precisely. In addition to linearizing in ε, we also use a perturbative expansion in λ, as argued for in Section 9.1.2. In fact, rather than working to first order in λ, we will work to second order in the Finslerian parameter, as certain important Finslerian effects only enter at second order, as we will see. We do neglect terms of combined order $\varepsilon\lambda^2$ and higher.

9.2.1 Null geodesics

As explained above, our goal is to find the unique two null geodesics that connect this point to the worldline of the observer. As long as the interferometer is sufficiently small in its spatial extent and the time window of observation is not too large, Proposition 8.3.7 guarantees that this can be done. The first important observation here is that the geodesics in a modified Randers gravitational wave spacetime with Finsler metric $F = \mathrm{sgn}(A)\alpha + |\beta|$ as in (9.8) coincide with the geodesics of the GR spacetime with metric A because the affine connection of F coincides with the Levi-Civita connection of A, by Corollary 5.2.2. For the derivation of the general form of geodesics, we may therefore assume that the geometry is given by A. Our point of departure will thus be the metric

$$\begin{aligned} &\mathrm{d}s^2 = \\ &-\mathrm{d}t^2 + (1+\varepsilon f_+(t-z))\mathrm{d}x^2 + (1-\varepsilon f_+(t-z))\mathrm{d}y^2 + 2\varepsilon f_\times(t-z)\mathrm{d}x\,\mathrm{d}y + \mathrm{d}z^2, \end{aligned} \tag{9.9}$$

where $\varepsilon \ll 1$, and we can essentially follow the corresponding calculation in GR [45] until the moment at which the modified null condition becomes important. Using the coordinates $u = (t-z)/\sqrt{2}$ and $v = (t+z)/\sqrt{2}$ the geodesic equations

to first order in ε can be written as

$$-\dot{u} =: p_v = \text{const}, \tag{9.10}$$

$$(1+\varepsilon f_+(u))\dot{x} + \varepsilon f_\times(u)\dot{y} =: p_x = \text{const}, \tag{9.11}$$

$$(1-\varepsilon f_+(u))\dot{y} + \varepsilon f_\times(u)\dot{x} =: p_y = \text{const}, \tag{9.12}$$

$$\ddot{v} + \tfrac{1}{2}\varepsilon\left(\dot{x}^2 - \dot{y}^2\right) f'_+(u) + \varepsilon f'_\times(u)\dot{x}\dot{y} = 0, \tag{9.13}$$

where (9.10)-(9.12) are obtained as first integrals using the fact that A is independent of the coordinates v, x, y. Here an overdot denotes a derivative with respect to the affine parameter σ. The first three equations can be rewritten to first order as $\dot{u} = -p_v$,

$$\dot{x} = (1-\varepsilon f_+(u))p_x - \varepsilon f_\times(u)p_y, \qquad \dot{y} = (1+\varepsilon f(u))p_y - \varepsilon f_\times(u)p_x, \tag{9.14}$$

and can be integrated (with respect to the affine parameter σ, chosen without loss of generality such that $\dot{u} = 1$) to

$$u = u_0 + \sigma, \tag{9.15}$$

$$x = x_0 + \sigma\left[\left(1 - \varepsilon\bar{f}_+(\sigma)\right)p_x - \varepsilon\bar{f}_\times(u)p_y\right], \tag{9.16}$$

$$y = y_0 + \sigma\left[\left(1 + \varepsilon\bar{f}_+(\sigma)\right)p_y - \varepsilon\bar{f}_\times(u)p_x\right], \tag{9.17}$$

where

$$\bar{f}_{+,\times}(\sigma) \equiv \frac{1}{\sigma}\int_0^\sigma f_{+,\times}(u_0+\sigma)\mathrm{d}\sigma \tag{9.18}$$

is the averaged value of $f_{+,\times}$. The equation (9.13) for v can be integrated to

$$\dot{v} = -\tilde{p}_{u0} - \frac{1}{2}\varepsilon\left(p_x^2 - p_y^2\right) f_+(u_0+\sigma) - \varepsilon f_\times(u_0+\sigma)p_x p_y \tag{9.19}$$

where $\tilde{p}_{u0} = p_{u0} - \frac{\varepsilon}{2}(p_x^2 - p_y^2)f(u_0) - \varepsilon f_\times(u_0)p_x p_y$ and where $p_u = -\dot{v}$ (not necessarily constant) and p_{u0} is its initial value at $\sigma = 0$. Integrating once again, we obtain

$$v = v_0 - \tilde{p}_{u0}\sigma - \frac{1}{2}\varepsilon(p_x^2 - p_y^2)\sigma\bar{f}_+(\sigma) - \varepsilon\bar{\sigma}f_\times(\sigma)p_x p_y. \tag{9.20}$$

Any geodesic emanating from a given point x_0^μ can thus be described by the following parameterized path for certain values of the constants p_x, p_y and $\tilde{p}_{u0}$:

$$u(\sigma) = u_0 + \sigma, \tag{9.21}$$
$$x(\sigma) = x_0 + \sigma\left[\left(1 - \varepsilon\bar{f}_+(\sigma)\right)p_x - \varepsilon\bar{f}_\times(\sigma)p_y\right], \tag{9.22}$$
$$y(\sigma) = y_0 + \sigma\left[\left(1 + \varepsilon\bar{f}_+(\sigma)\right)p_y - \varepsilon\bar{f}_\times(\sigma)p_x\right], \tag{9.23}$$
$$v(\sigma) = v_0 - \tilde{p}_{u0}\sigma - \tfrac{1}{2}\varepsilon(p_x^2 - p_y^2)\sigma\bar{f}_+(\sigma) - \varepsilon\sigma\bar{f}_\times(\sigma)p_x p_y. \tag{9.24}$$

We need to know the specific expression for *null* geodesics, however. The modified null condition or modified dispersion relation (MDR) for massless particles, $F = 0$, is the first place where the Finslerian character of the gravitational wave enters. According to Section 8.3.2, the condition $F = 0$ is equivalent to $A = -\beta^2$, i.e.

$$-2\dot{u}\dot{v} + (1 + \varepsilon f_+(u))\dot{x}^2 + (1 - \varepsilon f_+(u))\dot{y}^2 + 2\varepsilon f_\times(u)\dot{x}\dot{y} = -\lambda^2\dot{u}^2, \tag{9.25}$$

which, after substituting (9.21)-(9.24), becomes $2\tilde{p}_{u0} + p_x^2 + p_y^2 = -\lambda^2$. We may therefore eliminate $\tilde{p}_{u0}$ by directly substituting this null condition into the expression (9.24) for $v(\sigma)$. A generic null geodesic starting at (u_0, x_0, y_0, v_0) at $\sigma = 0$ can therefore be described by the following parameterized path,

$$u = u_0 + \sigma, \tag{9.26}$$
$$x = x_0 + \sigma\left[\left(1 - \varepsilon\bar{f}_+(\sigma)\right)p_x - \varepsilon\bar{f}_\times(u)p_y\right], \tag{9.27}$$
$$y = y_0 + \sigma\left[\left(1 + \varepsilon\bar{f}_+(\sigma)\right)p_y - \varepsilon\bar{f}_\times(u)p_x\right], \tag{9.28}$$
$$v = v_0 + \frac{\sigma}{2}(p_x^2 + p_y^2 + \lambda^2) - \frac{1}{2}\varepsilon(p_x^2 - p_y^2)\sigma\bar{f}_+(\sigma) - \varepsilon\sigma\bar{f}_\times(\sigma)p_x p_y. \tag{9.29}$$

Here we can make two important observations:

1. the effect due to the MDR or modified null condition enters at order λ^2;
2. in the limit $\lambda \to 0$ we recover the null geodesics of a standard gravitational wave in GR [45].

9.2.2 Radar distance

Next, we plug in the boundary conditions at the receiving point $(u_0 + \Delta u, x_0 + \Delta x, y_0 + \Delta y, v_0 + \Delta v)$. Note that $\sigma = \Delta u$ at that point, and hence from the

middle two equations (9.27) and (9.28) we infer that

$$p_x = \frac{\Delta x}{\Delta u}\left(1+\varepsilon \bar{f}_+(\Delta u)\right) + \varepsilon \bar{f}_\times(\Delta u)\frac{\Delta y}{\Delta u}, \tag{9.30}$$

$$p_y = \frac{\Delta y}{\Delta u}\left(1-\varepsilon \bar{f}_+(\Delta u)\right) + \varepsilon \bar{f}_\times(\Delta u)\frac{\Delta x}{\Delta u}, \tag{9.31}$$

Plugging this into the v equation yields

$$2\Delta u\Delta v = \Delta x^2(1+\varepsilon \bar{f}_+(\Delta u)) + \Delta y^2(1-\varepsilon \bar{f}_+(\Delta u)) + 2\varepsilon \bar{f}_\times \Delta x\Delta y + \lambda^2\Delta u^2, \tag{9.32}$$

or equivalently,

$$\left(1-\frac{\lambda^2}{2}\right)\Delta t^2 = \left(1+\varepsilon \bar{f}(\Delta u)\right)\Delta x^2 + \left(1-\varepsilon \bar{f}(\Delta u)\right)\Delta y^2 + 2\varepsilon \bar{f}_\times(\Delta u)\Delta x\Delta y + \left(1+\frac{\lambda^2}{2}\right)\Delta z^2 - \lambda^2\Delta z\Delta t, \tag{9.33}$$

where we have used that $-2\Delta u\Delta v = -\Delta t^2 + \Delta z^2$. This equation is solved to first order in ε and λ^2 (neglecting $\varepsilon\lambda^2$ terms) by[6]

$$\Delta t = \Delta \ell + \left(\frac{\Delta x^2 - \Delta y^2}{2\Delta \ell}\right)\varepsilon \bar{f}_+(\Delta u) + \left(\frac{\Delta x\Delta y}{\Delta \ell}\right)\varepsilon \bar{f}_\times(\Delta u) + \frac{1}{2}\left(\frac{\Delta x^2+\Delta y^2+2\Delta z^2}{2\Delta \ell} - \Delta z\right)\lambda^2, \tag{9.34}$$

where $\Delta \ell \equiv \sqrt{\Delta x^2+\Delta y^2+\Delta z^2}$ is the Euclidean spatial distance, which coincides with the radar distance in the case of flat Minkowski spacetime.

In principle, the RHS still depends on t via $\bar{f}(\Delta u)$, so (9.34) is not yet a closed formula for Δt. However, since $\bar{f}$ only appears multiplied with ε, and since we are only interested in the first order expression for Δt, any zeroth order expression for $\bar{f}$ suffices in this formula. We have

[6] In addition to this solution there is, formally, another solution to the equation. However, this other solution has the wrong zeroth order term, namely a negative one, which renders it physically irrelevant.

$$\bar{f}(\Delta u) = \frac{1}{\Delta u}\int_0^{\Delta u} f(u_0+\sigma)\mathrm{d}\sigma = \frac{\sqrt{2}}{\Delta t-\Delta z}\int_0^{(\Delta t-\Delta z)/\sqrt{2}} f(u_0+\sigma)\mathrm{d}\sigma$$
$$= \frac{\sqrt{2}}{\Delta \ell-\Delta z}\int_0^{(\Delta \ell-\Delta z)/\sqrt{2}} f(u_0+\sigma)\mathrm{d}\sigma + \mathcal{O}(\varepsilon) \tag{9.35}$$
$$= \frac{\sqrt{2}}{\Delta \ell-\Delta z}\int_0^{(\Delta \ell-\Delta z)/\sqrt{2}} f\left(\frac{1}{\sqrt{2}}(t_0-z_0)+\sigma\right)\mathrm{d}\sigma + \mathcal{O}(\varepsilon) \tag{9.36}$$

since $\Delta t = \Delta\ell + \mathcal{O}(\varepsilon)$. We introduce another symbol for this expression, namely

$$\bar{f}(\Delta\ell, \Delta z, t_0-z_0) \equiv \frac{\sqrt{2}}{\Delta \ell-\Delta z}\int_0^{(\Delta \ell-\Delta z)/\sqrt{2}} f\left(\frac{1}{\sqrt{2}}(t_0-z_0)+\sigma\right)\mathrm{d}\sigma, \tag{9.37}$$

where the explicit display of the arguments serves to remind us that $\bar{f}$ depends only on $\Delta\ell, \Delta z$ and the initial value of $t-z$. Since $\varepsilon\bar{f}(\Delta u) = \varepsilon\bar{f}(\Delta\ell, \Delta z, t_0 - z_0) + \mathcal{O}(\varepsilon^2)$, it follows that we can rewrite (9.34) to first order in ε and λ^2 as

$$\Delta t = \Delta\ell + \left(\frac{\Delta x^2-\Delta y^2}{2\Delta\ell}\right)\varepsilon\bar{f}_+(\Delta\ell, \Delta z, t_0-z_0)$$
$$+ \left(\frac{\Delta x\Delta y}{\Delta\ell}\right)\varepsilon\bar{f}_\times(\Delta\ell, \Delta z, t_0-z_0)$$
$$+ \frac{1}{2}\left(\frac{\Delta x^2+\Delta y^2+2\Delta z^2}{2\Delta\ell} - \Delta z\right)\lambda^2, \tag{9.38}$$

which is a closed expression for the elapsed coordinate time interval Δt for a light ray traveling over a certain spatial coordinate distance, in terms of the spatial coordinate separations and the initial value of $t-z$.

Now let us consider the complete trip, from x_0^μ to $x_0^\mu + \Delta x^\mu$ and 'back' to the original spatial location. The total coordinate time elapsed during this trip is the sum of the forward trip and the backward trip time intervals. Schematically:

$$\Delta t_{\mathrm{tot}} = \Delta t(\Delta x, \Delta y, \Delta z, t_0-z_0)$$
$$+ \Delta t(-\Delta x, -\Delta y, -\Delta z, t_0 + \Delta t - (z_0+\Delta z)), \tag{9.39}$$

since the spatial interval on the backward trip is simply minus the forward spatial interval, and since the 'initial' value of $t-z$ for the backward trip is just the final value $t_0 - z_0 + \Delta t - \Delta z$ corresponding to the forward trip. Plugging in (9.38)

yields

$$\Delta t_{\rm tot} = 2\Delta\ell + \varepsilon\left(\frac{\Delta x^2 - \Delta y^2}{2\Delta\ell}\right)\bar{f}_{+,\rm tot} + \varepsilon\left(\frac{\Delta x\Delta y}{\Delta\ell}\right)\bar{f}_{\times,\rm tot} + \frac{1}{2}\lambda^2\left(\frac{\Delta x^2 + \Delta y^2 + 2\Delta z^2}{\Delta\ell}\right), \tag{9.40}$$

where $\bar{f}_{+,\rm tot} = \bar{f}_{+,\rm forward} + \bar{f}_{+,\rm backward}$ and similarly for the $\times$-polarization, in terms of the forward and backward averaged amplitudes, respectively, given by

$$\begin{aligned}
\bar{f}_{+,\times,\rm forward} &= \bar{f}_{+\times,}(\Delta\ell, \Delta z, t_0 - z_0) \\
&= \frac{\sqrt{2}}{\Delta\ell - \Delta z}\int_0^{(\Delta\ell-\Delta z)/\sqrt{2}} f_{+,\times}\left(\frac{1}{\sqrt{2}}(t_0 - z_0) + \sigma\right) \mathrm{d}\sigma, \qquad (9.41) \\
\bar{f}_{+,\times,\rm backward} &= \bar{f}_{+,\times}(\Delta\ell, -\Delta z, t_0 - z_0 + \Delta t - \Delta z) \qquad (9.42) \\
&= \frac{\sqrt{2}}{\Delta\ell + \Delta z}\int_0^{(\Delta\ell+\Delta z)/\sqrt{2}} f_{+,\times}\left(\frac{1}{\sqrt{2}}(t_0 + \Delta\ell - z_0 - \Delta z) + \sigma\right) \mathrm{d}\sigma.
\end{aligned}$$

In the last expression we have replaced Δt by $\Delta\ell$ in the argument of $f_{+,\times}$, because to zeroth order this makes no difference, and only the zeroth order expression for $\bar{f}_{+,\rm backward}$ is relevant because $\bar{f}_{+,\rm backward}$ always appears multiplied by ε in the expressions we care about, like $\Delta t_{\rm tot}$.

Equation (9.40) gives the total coordinate time elapsed during the trip of the light ray forward and back. The next step in the calculation of the radar distance $R = \Delta\tau/2$ is to convert the coordinate time interval into the proper time interval measured by the stationary observer local to the emission and reception of the light ray. This is where a second Finslerian effect enters. For such an observer we have $x = y = z = const$ and hence the 4-velocity is given by $(\dot{t}, 0, 0, 0)$, where we will assume without loss of generality that $\dot{t} > 0$. The proper time measured by an observer is given by the Finslerian length along its worldline $\Delta\tau = -\int F\,\mathrm{d}\sigma$. If we use $\sigma = \tau$ as our curve parameter, differentiating with respect to it shows that the tangent vector to the curve should be normalized, satisfying $F = -1$. This is the Finsler equivalent of the fact that in GR the worldline of a particle parameterized by proper time should always satisfy $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = -1$ (or $+1$, depending on the signature convention). In the case of our observer, the condition

becomes

$$F = \mathrm{sgn}(A)\alpha + |\beta| = \mathrm{sgn}(-\dot{t}^2)\sqrt{|\dot{t}^2|} + \frac{|\lambda\dot{t}|}{\sqrt{2}} = -|\dot{t}| + \frac{|\lambda\dot{t}|}{\sqrt{2}} \tag{9.43}$$

$$= \left(-1 + \frac{\lambda}{\sqrt{2}}\right)\dot{t} \stackrel{!}{=} -1, \tag{9.44}$$

where we recall that $\lambda > 0$. From this, it follows that

$$\Delta\tau = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta t_{\mathrm{tot}} \tag{9.45}$$

along the worldline of the stationary observer. Plugging in (9.40) and (9.45) into $R = \Delta\tau/2$ we conclude that, to first order in ε and second order in λ, the radar distance is given by

$$R = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta\ell + \left(1 - \frac{\lambda}{\sqrt{2}}\right)\left(\frac{\Delta x^2 - \Delta y^2}{4\Delta\ell}\right)\varepsilon\bar{f}_{+,\mathrm{tot}} \tag{9.46}$$

$$+ \left(1 - \frac{\lambda}{\sqrt{2}}\right)\left(\frac{\Delta x\Delta y}{2\Delta\ell}\right)\varepsilon\bar{f}_{\times,\mathrm{tot}} + \frac{\lambda^2}{4}\left(\Delta\ell + \frac{\Delta z^2}{\Delta\ell}\right).$$

Equation (9.46) expresses the radar distance as a function of the spatial coordinate distances and the initial value of $t - z$ (the latter enters the expression via $\bar{f}_{+,\times,\mathrm{tot}}$). In the limit $\lambda \to 0$ we recover the expression for the radar distance in the case of a standard gravitational wave in GR [45]:

$$R = \Delta\ell + \varepsilon\left(\frac{\Delta x^2 - \Delta y^2}{4\Delta\ell}\right)\bar{f}_{+,\mathrm{tot}} + \varepsilon\left(\frac{\Delta x\Delta y}{2\Delta\ell}\right)\bar{f}_{\times,\mathrm{tot}} + \mathcal{O}(\varepsilon^2). \tag{9.47}$$

Before we move on, let us summarize in what ways the Finslerian parameter λ has entered our analysis so far:

1. The null geodesics are altered due to the fact the Finsler metric induces a modified null condition or MDR. As a result, it takes a *larger* coordinate time interval for a light ray to travel a given spatial coordinate distance. This effect works in all spatial directions, even the direction parallel to the propagation direction of the light ray (although the effect is somewhat different in this direction). This effect enters at order λ^2.

2. The ratio of proper time and coordinate time is altered with the result that *less proper time is experienced per unit coordinate time*. This effect enters at order λ.

There is, however, a third way in which the parameter enters. Namely in the relation between the coordinate distance and radar distance *in the absence of the wave*. For a gravitational wave in GR, these conveniently coincide; in the case of our Randers waves, they do not. The expression (9.46) for the radar distance derived above refers merely to coordinates. In order to make sense of the result, we need to express the RHS in terms of measurable quantities, like the radar distances in the various directions in the absence of the wave. Employing (9.46), we write

$$\Delta X = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta x + \frac{\lambda^2}{4}\Delta x, \tag{9.48}$$

$$\Delta Y = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta y + \frac{\lambda^2}{4}\Delta y, \tag{9.49}$$

$$\Delta Z = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta z + \frac{\lambda^2}{2}\Delta z, \tag{9.50}$$

for the radar distance in the x, y and z direction *in the absence of the wave*, respectively, and

$$R_0 = \left(1 - \frac{\lambda}{\sqrt{2}}\right)\Delta\ell + \frac{\lambda^2}{4}\left(\Delta\ell + \frac{\Delta z^2}{\Delta\ell}\right), \tag{9.51}$$

for the radar distance (9.46) in the relevant direction *in the absence of the wave*. Eliminating the coordinate distances in favor of the physical radar distances by virtue of the inverse transformations

$$\Delta x = \Delta X\left(1 + \frac{\lambda}{\sqrt{2}} + \frac{\lambda^2}{4}\right), \tag{9.52}$$

$$\Delta y = \Delta Y\left(1 + \frac{\lambda}{\sqrt{2}} + \frac{\lambda^2}{4}\right), \tag{9.53}$$

$$\Delta z = \Delta Z\left(1 + \frac{\lambda}{\sqrt{2}}\right), \tag{9.54}$$

$$\Delta\ell = R_0\left(1 + \frac{\lambda}{\sqrt{2}} + \frac{3}{4}\lambda^2\right) - \frac{\Delta z^2}{4R_0}\lambda^2, \tag{9.55}$$

$$= R_0\left(1 + \frac{\lambda}{\sqrt{2}} + \frac{\lambda^2}{4}\right) - \frac{\Delta Z^2}{4R_0}\lambda^2, \tag{9.56}$$

valid to second order in λ,we obtain our final expression for the radar distance

in the presence of the wave:

$$R = R_0 + \varepsilon \left(\frac{\Delta X^2 - \Delta Y^2}{4R_0} \right) \bar{f}_{+,\mathrm{tot}} + \varepsilon \left(\frac{\Delta X \Delta Y}{2R_0} \right) \bar{f}_{\times,\mathrm{tot}} + \mathcal{O}(\varepsilon^2, \lambda^3, \varepsilon\lambda^2). \tag{9.57}$$

This is a remarkable result. By expressing the radar distance in terms of the physical observables $\Delta X, \Delta Y$ and R_0 rather than merely coordinate separations, all dependence on λ has disappeared to the desired order and the expression is identical to its GR counterpart (9.47)! We must conclude, therefore, that the effect of a Randers gravitational wave on an interferometer is virtually indistinguishable from that of a conventional GR gravitational wave.

9.3 Concluding remarks

By obtaining an explicit expression for the radar distance of our Finslerian gravitational waves, we have shown that interferometer experiments are unable to distinguish such gravitational waves from general relativistic ones. This invites some discussion. On the one hand, the result is a bit disappointing since it indicates that such measurements cannot be used to test the Finslerian character of our universe. On the other hand, it also means that current gravitational wave measurements are all compatible with the idea that our spacetime has a Finslerian nature. Further research (theoretical as well as experimental) is necessary to obtain a definite answer to the question of whether Finslerian gravitational waves exist, or more generally, the question of whether our universe has a Finslerian nature.

In this context, our radar distance result may be viewed not in the least as a proof of concept. It shows how clear physical predictions can be extracted from an *a priori* abstract Finslerian model of spacetime. In this particular case, the predictions coincide with those of GR, but it would be highly worthwhile to repeat such an analysis in other cases of interest.

It is important to point out that Finslerian effects may also play a role in the *generation* of gravitational waves during, say, a black hole merger event. This would lead to a Finslerian correction to the waveform and this *could* be measured in interferometer experiments, at least in principle; especially since our results show how the waveform correlates with the experimental results—namely in exactly the same way in which the waveform of a general relativistic gravitational wave does. In order to be able to investigate this, however, Finslerian black hole solutions need to be better understood. A start in this direction has been made

in [121, 102], and further research (in part by the author) is actively ongoing.

Let us also point out some of the limitations of our investigation. First of all, it is by no means expected that the Finslerian gravitational waves discussed here should be the only possible ones. Although being much larger than even the complete class of *all* Lorentzian (i.e. non-Finslerian) geometries, the class of (α, β)-metrics of Berwald type, to which we have restricted our analysis, is still quite restrictive in the large scheme of (Finsler geometric) things.

Moreover, even within the class of (α, β)-metrics, our analysis is only valid for those metrics that can be regarded as 'close' to a Lorentzian metric, such that they can be approximated by Randers metrics. So even though there is no observable difference between the Finslerian gravitational waves discussed here and their GR counterparts, there might very well be more general types of Finslerian gravitational waves that *could* be distinguished observationally from general relativistic ones by means of interferometer experiments.

Furthermore, radar distance experiments are by no means the only way of probing our spacetime geometry. It might be possible to detect the Finslerian character of spacetime in some other way. We have not explored such possibilities here, but will certainly do so in the future.

Finally, we have assumed in our calculations that the amplitude of the gravitational waves as well as the Finslerian deviation from general relativity are sufficiently small such that a perturbative approach to first order in the former and second order in the latter is valid. It would be of interest to repeat the calculation to higher order in perturbation theory. We expect that this would in principle be a straightforward, yet possibly tedious, exercise.

CHAPTER 10

An Exact Cosmological Unicorn Solution

Berwald and Landsberg spacetimes (see Chapter 4) may be thought of as incrementally non–pseudo-Riemannian. Or in physics terms: incrementally less reminiscent of GR. Every Berwald spacetime is also Landsberg, but provided one adheres to the most strict definition of a Finsler metric (more strict than ours) it has been a long-standing open question whether or not the opposite is true. In fact, Matsumoto stated in 2003 that this question represents the next frontier of Finsler geometry [46], and as a token of their elusivity, Bao [46] has called these non-Berwaldian Landsberg spaces *'[. . .] unicorns, by analogy with those mythical single-horned horse-like creatures for which no confirmed sighting is available.'* Since 2006 some examples of unicorns have been obtained by Asanov [47], Shen [48] and Elgendi [49] by relaxing the definition of a Finsler space slightly, resulting in a definition like ours. Even such examples of so-called y-local unicorns are still exceedingly rare.

In this chapter, which is based on [H6], we present a new exact vacuum solution to Pfeifer and Wohlfarth's field equation, which is precisely such a unicorn. It falls into one of the classes introduced by Elgendi. Interestingly we find that these solutions have a physically viable light cone structure, even though in some cases the signature is not Lorentzian but positive definite. In fact, the Finslerian light cone turns out to be equivalent to that of the flat Minkowski metric, even in the cases with the signature anomaly!

Furthermore, we find a natural cosmological interpretation of one of our solutions and a promising analogy with classical Friedmann-Lemaître-Robertson-

Walker (FLRW) cosmology. In particular, this solution has cosmological symmetry, i.e. it is spatially homogeneous and isotropic, and it is additionally conformally flat, with conformal factor depending only on the timelike coordinate. We show that, just as in classical FLRW cosmology, this conformal factor can be interpreted as the scale factor of the universe; we compute this scale factor as a function of cosmological time, and we show that it corresponds to a linearly expanding (or contracting) Finslerian universe.

10.1 A modification of Elgendi's unicorn metrics

10.1.1 Elgendi's class of unicorns

Elgendi recently introduced a class of unicorns [49] with Finsler metric given by

$$F = \left(a\beta + \sqrt{\alpha^2 - \beta^2}\right) e^{\frac{a\beta}{a\beta+\sqrt{\alpha^2-\beta^2}}}, \tag{10.1}$$

in terms of a real, nonvanishing constant a (not to be confused with the metric $a_{\mu\nu}$ this time) and

$$\alpha = f(x^0)\sqrt{(y^0)^2 + \phi(\hat{y})}, \qquad \beta = f(x^0)y^0, \tag{10.2}$$

where f is a positive real-valued function and $\phi(\hat{y}) = \phi_{ij}\hat{y}^i\hat{y}^j = \phi_{ij}y^iy^j$ is a nondegenerate quadratic form on the space spanned by $\hat{y} = (y^1, y^2, y^3)$, with constant, symmetric coefficients ϕ_{ij}. As before, Latin indices $i, j, \ldots$ will run over $1, 2, 3$, whereas Greek induces $\mu, \nu, \ldots$ will run over $0, 1, 2, 3$. The geodesic spray of F is given by

$$G^0 = \left(\frac{2f(x^0)^2(y^0)^2 - \alpha^2}{f(x^0)^2} + \frac{a^2-1}{a^2}\frac{\alpha^2-\beta^2}{f(x^0)^2}\right)\frac{f'(x^0)}{f(x^0)} \tag{10.3}$$

$$G^i = Py^i, \tag{10.4}$$

where

$$P = 2\left(y^0 + \frac{1}{af(x^0)}\sqrt{\alpha^2-\beta^2}\right)\frac{f'(x^0)}{f(x^0)}, \tag{10.5}$$

and the Landsberg tensor vanishes identically. This shows that these metrics are indeed Landsberg, but not Berwald, since the i-components of the spray are not

quadratic[1] in y. Note that our G^k is twice the G^k in Elgendi's paper [49], due to a difference in convention. Explicitly then, Elgendi's unicorns have the form

$$F = f(x^0)\left(y^0 + \sqrt{\phi(\hat{y})}\right) e^{\frac{y^0}{y^0+\sqrt{\phi(\hat{y})}}}. \tag{10.6}$$

where we have absorbed the constant a into a redefinition of x^0. (And so from here onward there is no possibility for confusion of the constant a with the metric $a_{\mu\nu}$ anymore.) For our purposes, we will modify this expression slightly, though, for similar reasons as the ones that motivated our definition of the modified Randers metric in Section 8.3.

10.1.2 The modified unicorn metric

The expression (10.6) defining the unicorn metric is only well-defined whenever $\phi(\hat{y}) \geq 0$. If ϕ is positive definite, this is automatically the case, but in other signatures this is not true in general. In order to extend the domain of definition of F, an obvious first approach would be to replace ϕ by its absolute value, $|\phi|$, leading to

$$F = f(x^0)\left(y^0 + \sqrt{|\phi(\hat{y})|}\right) e^{\frac{y^0}{y^0+\sqrt{|\phi(\hat{y})|}}}. \tag{10.7}$$

From the physical point of view, this is still not completely satisfactory, however. This can be seen by considering the light cone corresponding to such a Finsler metric, given by $F = 0$. Indeed, barring some potential problems with the exponent to which we will come back later, the light cone would be given by those vectors satisfying $y^0 = -\sqrt{|\phi|}$, which would imply that the light cone is contained entirely within the half-space $y^0 < 0$. Interpreting y^0 for the moment as a time direction, this would have the result that light rays can only propagate 'backward in time' (with regards to their parameterization). If on the other hand, y^0 is a spacelike coordinate, the analogous statement would be that light cannot propagate in the spatial y^0-direction (in contrast to the $-y^0$-direction). One could perhaps argue that these distinctions only pertain to the light rays as *parameterized* curves and hence are not necessarily physical, but even as a mathematical artifact, it seems odd.

[1] Indeed, if each G^i were quadratic this would imply that $\sqrt{\alpha^2 - \beta^2} = (c_i y^i)^2$ for certain (x-dependent) coefficients c_i. Rewriting gives $\alpha^2 = \beta^2 + (c_i y^i)^2$ and taking two y-derivatives of this (in the open set $A > 0$ for concreteness) shows that $a_{ij} = b_i b_j + c_i c_i$. That implies that a_{ij} has at most rank 2, which is in contradiction with its definition as a 4-dimensional Lorentzian metric, which should have rank 4.

This situation is reminiscent of the similar situation for Randers metrics discussed extensively in Section 8.3. The solution in that case was to replace the Randers metric with the modified Randers metric (8.24). Here we will perform a similar modification of the unicorn metric based on the same ideas.

Our starting point will be the following Finsler metric:

$$F_0 = f(x^0)\left(|y^0| + \operatorname{sgn}(\phi)\sqrt{|\phi|}\right) e^{\frac{|y^0|}{|y^0|+\operatorname{sgn}(\phi)\sqrt{|\phi|}}}, \tag{10.8}$$

from which we define the *modified unicorn metric* as:

$$F = \begin{cases} F_0 & \text{if} \quad |y^0| + \operatorname{sgn}(\phi)\sqrt{|\phi|} \neq 0 \\ 0 & \text{if} \quad |y^0| + \operatorname{sgn}(\phi)\sqrt{|\phi|} = 0 \end{cases}. \tag{10.9}$$

It will be confirmed in Section 10.2 that this still defines a non-Berwaldian Landsberg metric, justifying the term *unicorn*.

The case distinction is necessary since the metric (10.8) is ill-defined at vectors satisfying $|y^0|+\operatorname{sgn}(\phi)\sqrt{|\phi|} = 0$, because of the division by the same number in the exponent. F_0 does not have a well-defined limit to such vectors either, because the exponent does not stay negative in such a limit. This issue was already present for Elgendi's unicorn (10.6), yet from a purely mathematical point of view, this is not necessarily a problem, as one can simply opt to exclude this set of vectors from the domain of F. From a physics perspective, however, we want to interpret the set $F = 0$ as the possible propagation directions of light, so it should not be empty.

As we will see below, our definition of F as in (10.9) ensures the existence of a physically viable light cone while preserving also the unicorn property. This why we prefer it in the physical context, even though it leads to a discontinuity at the light cone as discussed in more detail below in Section 10.1.4.

10.1.3 Causal structure

First, we observe that regardless of the exact form or signature of the quadratic form ϕ, our modified unicorn metrics have a light cone structure $F = 0$ that is equivalent to that of a pseudo-Riemannian metric.

Proposition 10.1.1. *The light cone of the modified unicorn metric* (10.9) *is given by*

$$\left(y^0\right)^2 + \phi = 0. \tag{10.10}$$

Proof. The result follows from the following sequence of equivalences.

$$F = 0 \Leftrightarrow \operatorname{sgn}(\phi)\sqrt{|\phi|} + |y^0| = 0 \tag{10.11}$$

$$\Leftrightarrow \operatorname{sgn}(\phi)\sqrt{|\phi|} = -|y^0| \tag{10.12}$$

$$\Leftrightarrow |\phi| = \left(y^0\right)^2 \quad \text{and} \quad \phi \leq 0 \tag{10.13}$$

$$\Leftrightarrow \phi = -\left(y^0\right)^2 \tag{10.14}$$

$$\Leftrightarrow \phi + \left(y^0\right)^2 = 0. \tag{10.15}$$

□

Depending on the signature of ϕ, we can make a more precise statement. In the following corollary, the ordering of the plusses and minuses in expressions like $(+,+,+,-)$ and $(-,+,+,+)$ indicates which of the coordinates—in this case, x^3 and x^0, respectively—is the timelike one.

Corollary 10.1.2. *Let F be the modified unicorn metric (10.9) corresponding to some nondegenerate quadratic form ϕ_{ij}. Then the following holds:*

- *For ϕ_{ij} of signature $(+,+,-)$ the light cone structure of F is identical to that of the $(+,+,+,-)$ Minkowski metric.*
- *If ϕ_{ij} is negative definite, i.e. of signature $(-,-,-)$ then the light cone of F is identical to that of the $(-,+,+,+)$ Minkowski metric.*
- *If ϕ_{ij} is positive definite, i.e. of signature $(+,+,+)$ then the light cone of F is given by the zero vector, $y = 0$.*
- *For ϕ_{ij} of signature $(-,-,+)$ the light cone structure of F is identical to that of a pseudo-Riemannian metric of signature $(+,-,-,+)$.*

This singles out the $(+,+,-)$ and $(-,-,-)$ signatures of ϕ_{ij} as the ones that are physically of interest, and we will mostly restrict our attention to these two cases in what follows. Given the fact that the light cone structure in these cases is so simple, it is natural—as discussed several times before—to interpret the interior of the future– and past-pointing light cone as the cone of future– and past-pointing timelike directions, respectively. In the $(+,+,-)$ case, this leads to the interpretation of the coordinate x^3 as the timelike coordinate, with the timelike cone given by $(y^3)^2 > (y^0)^2 + (y^1)^2 + (y^2)^2$, while in the $(-,-,-)$ case,

x^0 would be the timelike coordinate, with the timelike cone given by $(y^0)^2 > (y^1)^2 + (y^2)^2 + (y^3)^2$.

In analogy to the discussion in Section 8.3.2 it also follows, in the same way that it did for modified Randers metrics, that the sign of $F(x, y)$ is correlated with the causal character of y. In the current scenario, however, the correlation depends also on the signature of ϕ. The details are provided by the following proposition.

Proposition 10.1.3 (Modified unicorn causal structure)**.**

- *If ϕ has signature $(-,-,-)$ then*

$$\begin{aligned} y \text{ is timelike} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} > 0 \quad\Leftrightarrow\quad F > 0 \\ y \text{ is spacelike} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} < 0 \quad\Leftrightarrow\quad F < 0 \\ y \text{ is null} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} = 0 \quad\Leftrightarrow\quad F = 0 \end{aligned}$$

- *If ϕ has signature $(+,+,-)$ then*

$$\begin{aligned} y \text{ is timelike} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} < 0 \quad\Leftrightarrow\quad F < 0 \\ y \text{ is spacelike} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} > 0 \quad\Leftrightarrow\quad F > 0 \\ y \text{ is null} &\quad\Leftrightarrow\quad |y^0| + sgn(\phi)\sqrt{|\phi|} = 0 \quad\Leftrightarrow\quad F = 0 \end{aligned}$$

Proof. We start with the $F < 0$ equivalences. Since $f(x^0)$ is positive by assumption, we have

$$\begin{aligned} F < 0 &\Leftrightarrow \operatorname{sgn}(\phi)\sqrt{|\phi|} + |y^0| < 0 && (10.16)\\ &\Leftrightarrow \operatorname{sgn}(\phi)\sqrt{|\phi|} < -|y^0| && (10.17)\\ &\Leftrightarrow |\phi| > (y^0)^2 \quad \text{and} \quad \phi \leq 0 && (10.18)\\ &\Leftrightarrow -\phi > (y^0)^2 && (10.19)\\ &\Leftrightarrow (y^0)^2 + \phi < 0. && (10.20) \end{aligned}$$

In the case with signature $(-,-,-)$ this inequality reads $(y^0)^2 - (y^1)^2 - (y^2)^2 - (y^3)^2 < 0$, which says precisely that y lies in the exterior of the light cone $(y^0)^2 = (y^1)^2 + (y^2)^2 + (y^3)^2$, i.e. y is spacelike. And in the $(+,+,-)$ case the

inequality reads $(y^0)^2+(y^1)^2+(y^2)^2-(y^3)^2<0$, which says precisely that y lies in the interior of the light cone $(y^0)^2+(y^1)^2+(y^2)^2=(y^3)^2$, i.e. y is timelike. This completes the proof of the $F<0$ equivalences. The $F=0$ equivalences follow trivially from the definition of a null vector and Proposition 10.1.1. Since $F>0$ if and only if neither of the previous two cases $F<0$ or $F=0$ is realized, the remaining $F>0$ equivalences follow automatically. □

10.1.4 Signature and regularity

Next, we investigate the signature of our modified unicorn metrics. When ϕ has Lorentzian signature, we denote by $\mathcal{S}(\phi)$ the set of all $\hat{y}=(y^1,y^2,y^3)$ that are ϕ-spacelike and by $\mathcal{T}(\phi)$ the set of all $\hat{y}$ that are ϕ-timelike.

Proposition 10.1.4. *Consider a modified unicorn metric F as in (10.9).*

- *If ϕ is positive definite or negative definite then $g_{\mu\nu}$ is positive definite on its entire domain of definition.*
- *If ϕ is Lorentzian then $g_{\mu\nu}$ has Lorentzian signature $(+,+,+,-)$ for all $y\in\mathbb{R}\times\mathcal{S}(\phi)$ at which it is defined and signature $(+,+,-,-)$ for all $y\in\mathbb{R}\times\mathcal{T}(\phi)$ at which it is defined.*

Proof. We may choose coordinates such that $\phi=\phi(\hat{y})=\varepsilon_1(y^1)^2+\varepsilon_2(y^2)^2+\varepsilon_3(y^3)^2$. Since the spacetime dimension is fixed to 4 the calculation of the determinant of the fundamental tensor is in principle a straightforward exercise whenever F is sufficiently differentiable. It is given by

$$\det g=\operatorname{sgn}(\phi)\varepsilon_1\varepsilon_2\varepsilon_3 f(x^0)^8\exp\left(\frac{8|y^0|}{|y^0|+\operatorname{sgn}(\phi)\sqrt{|\phi|}}\right). \tag{10.21}$$

The determinant already gives us a pretty good idea of what the possible signatures of $g_{\mu\nu}$ can be. In particular, since $g_{\mu\nu}$ is a four-dimensional matrix, it has Lorentzian signature if and only if its determinant is negative.

a) If ϕ_{ij} is positive definite then all ε_i and $\operatorname{sgn}(\phi(\hat{y}))$ are positive, and hence $\det g$ is positive.

b) If ϕ_{ij} is negative definite then all ε_i and $\operatorname{sgn}(\phi(\hat{y}))$ are negative, and hence $\det g$ is positive.

c) If ϕ_{ij} has Lorentzian signature $(+,+,-)$ then $\det g$ is negative whenever $\operatorname{sgn}(\phi(\hat{y})) > 0$, i.e. on $\mathcal{S}(\phi)$, and $\det g$ is positive whenever $\operatorname{sgn}(\phi(\hat{y})) < 0$, i.e. on $\mathcal{T}(\phi)$.

d) If ϕ_{ij} has Lorentzian signature $(-,-,+)$; then $\det g$ is negative whenever $\operatorname{sgn}(\phi(\hat{y})) < 0$, i.e. on $\mathcal{S}(\phi)$, and $\det g$ is positive whenever $\operatorname{sgn}(\phi(\hat{y})) > 0$, i.e. on $\mathcal{T}(\phi)$.

This already shows that $g_{\mu\nu}$ is Lorentzian if and only if ϕ is Lorentzian and $y \in \mathbb{R} \times \mathcal{S}(\phi)$. But the sign of the determinant does not suffice to determine whether this signature is mostly plus or mostly minus. Similarly, it does not tell us much about the signature of $g_{\mu\nu}$ when ϕ is positive or negative definite. In order to find out, we distinguish the following three cases.

Case 1: ϕ Lorentzian and $y \in \mathbb{R} \times \mathcal{S}(\phi)$
We first consider the case that ϕ is Lorentzian. Without loss of generality, we set $\phi(\hat{y}) = \varepsilon_1(y^1)^2 + \varepsilon_1(y^2)^2 - \varepsilon_1(y^3)^2$, where $\varepsilon_1 = \pm 1$ selects if we are in case c) or d).

Now note that given a vector $y \in T_xM$ which is ϕ-spacelike, it follows from the symmetries of the Finsler metric and in particular from the 3-dimensional Lorentz symmetry of ϕ that we may always change coordinates, without changing the form of ϕ (and F), such that $y^2 = y^3 = 0$.

For any choice of ε_1, we find by direct calculation, using that $\varepsilon_1^2 = 1$ and $|\varepsilon_1| = 1$, that $g_{\mu\nu}$ is of the form

$$g_{\mu\nu} = e^{\frac{2|y^0|}{|y^0|+\varepsilon_1|y^1|}} f(x^0)^2 \begin{pmatrix} M & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{10.22}$$

where M is an (ε_1-dependent) 2×2 matrix, whose explicit form we omit. Its relevant properties are that

$$\det M = 1, \qquad \operatorname{Tr} M = 2 + \frac{4(y^1)^2}{(|y^0| + \varepsilon_1|y^1|)^2}. \tag{10.23}$$

Since the determinant and trace are both positive it follows that M is positive definite. Hence we conclude that $g_{\mu\nu}$ is of the mostly plus type $(+,+,+,-)$.

Case 2: ϕ Lorentzian and $y \in \mathbb{R} \times \mathcal{T}(\phi)$
In this case we may WLOG choose coordinates such that $\phi(\hat{y}) = \varepsilon_1(y^1)^2 - \varepsilon_1(y^2)^2 - \varepsilon_1(y^3)^2$, where $\varepsilon_1 = \pm 1$, and such that $y^2 = y^3 = 0$. Again, by direct

calculation we find that $g_{\mu\nu}$ is now of the form

$$g_{\mu\nu} = e^{\frac{2|y^0|}{|y0|+\varepsilon_1|y1|}} f(x^0)^2 \begin{pmatrix} M & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{10.24}$$

where M is again a 2×2 matrix with the properties (10.23), so that M is again positive definite. Hence we conclude that in this case, $g_{\mu\nu}$ has signature $(+,+,-,-)$.

Case 3: ϕ positive or negative definite
In this case we may WLOG choose coordinates such that $\phi(\hat{y}) = \varepsilon_1(y^1)^2 + \varepsilon_1(y^2)^2 + \varepsilon_1(y^3)^2$, where $\varepsilon_1 = \pm 1$, and such that, for any given $y \in T_xM$, we have $y^2 = y^3 = 0$. In this case $g_{\mu\nu}$ is of the form

$$g_{\mu\nu} = e^{\frac{2|y^0|}{|y0|+\varepsilon_1|y1|}} f(x^0)^2 \begin{pmatrix} M & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{10.25}$$

where M is again a 2×2 matrix with the properties (10.23), so that M is again positive definite. Hence we conclude that $g_{\mu\nu}$ is positive definite. □

Let us point out some interesting features of the two physically reasonable scenarios we identified below Corollary 10.1.2 as a result of their satisfactory light cone structure, i.e. ϕ having $(-,-,-)$ or $(+,+,-)$ signature. In both cases the light cone is equivalent to that of Minkowski space, but surprisingly, in its interior—the timelike cone—the signature of g is never Lorentzian. Indeed, if ϕ is negative definite then Proposition 10.1.4 shows that g is positive definite everywhere, in particular inside the timelike cone. Similarly, if ϕ has signature $(+,+,-)$ then, with x^3 representing the timelike direction as pointed out right below Corollary 10.1.2, the timelike cone is given by $(y^3)^2 > (y^0)^2 + (y^1)^2 + (y^2)^2$. Any vector in this cone thus satisfies, in particular, the inequality $(y^3)^2 > (y^1)^2 + (y^2)$, which says precisely that $\hat{y} \in \mathcal{T}(\phi)$. It then follows from Proposition 10.1.4 that g has signature $(+,+,-,-)$ here and hence everywhere within the timelike cone.

In various alternative, more stringent definitions of Finsler spacetimes [97, 122], one requires the existence of a nonempty cone with certain properties on which the fundamental tensor has Lorentzian signature, in order to guarantee (among other things) the existence of a physically viable light cone structure. What we have just established here, however, is that there exist Finsler geometries that do have a perfectly viable light cone structure even while not having

Lorentzian signature *anywhere*. This is an interesting new observation about Finsler geometry in its own right: apparently, even in positive definite signature, a light cone structure may arise due to irregularities of the Finsler metric. The most notable irregularity in the present case is located at null vectors, i.e. vectors satisfying $|y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|} = 0$. Indeed, symbolically, we have

$$\lim_{\substack{|y^0|+\mathrm{sgn}(\phi)\sqrt{|\phi|}\to 0 \\ |y^0|+\mathrm{sgn}(\phi)\sqrt{|\phi|}<0}} F = 0. \tag{10.26}$$

whereas

$$\lim_{\substack{|y^0|+\mathrm{sgn}(\phi)\sqrt{|\phi|}\to 0 \\ |y^0|+\mathrm{sgn}(\phi)\sqrt{|\phi|}>0}} F = \infty\,. \tag{10.27}$$

Taking into account Proposition 10.1.3 this means that in the scenario where ϕ has signature $(+,+,-)$, the Finsler metric F extends continuously to 0 if one tends to the light cone from inside the timelike cone. The transition from the spacelike directions to the light cone is then discontinuous, however. In the $(-,-,-)$ scenario the situation is exactly the opposite. F then extends continuously to 0 if one tends to the light cone from the spacelike directions, whereas the transition from the timelike directions to the light cone is discontinuous.

10.2 Exact solutions

Our next aim is to determine the form of the function $f(x^0)$—the only physical degree of freedom—in the modified unicorn metric (10.9) by solving the Finsler gravity field equation. Since the latter simplifies tremendously for Landsberg spaces, let us start by confirming that the modified unicorn is still of this type. The geodesic spray of F is given explicitly by

$$G^0 = \left((y^0)^2 - |\phi|\right)\frac{f'(x^0)}{f(x^0)} \tag{10.28}$$

$$G^i = Py^i, \qquad i = 1,2,3 \tag{10.29}$$

where

$$P = 2\left(|y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|}\right)\mathrm{sgn}(y^0)\frac{f'(x^0)}{f(x^0)}\,, \tag{10.30}$$

In analogy with Elgendi's original unicorn metrics (10.6), this geodesic spray is not quadratic in y, so the modified unicorn metrics (10.9) are not Berwald. Let us now check that the Landsberg tensor still vanishes, justifying the name (modified) *unicorn*. To this end we employ the definition (4.5) of the Landsberg tensor, $S_{\mu\nu\rho} = -\frac{1}{4} y_\sigma \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\rho G^\sigma$, and note that only the G^i terms ($i = 1, 2, 3$) can give a nontrivial contribution, since G^0 is quadratic. For these terms, we compute that

$$\bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma G^i = \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma \sqrt{|\phi|} y^i + 3\delta^i{}_{(\mu} \bar{\partial}_\nu \bar{\partial}_{\sigma)} \sqrt{|\phi|}, \tag{10.31}$$

where the round brackets around the indices denote symmetrization. To find the Landsberg tensor we need to contract this with $y_i = g_{i\mu} y^\mu = \frac{1}{2} \bar{\partial}_i F^2$, and it can be checked in a straightforward way that the latter can be written as some function times $\bar{\partial}_i \sqrt{|\phi|}$. It thus suffices to show that $\bar{\partial}_i \sqrt{|\phi|} \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma G^i = 0$. Indeed, whenever $\mu \neq 0$, it follows from the homogeneity of $\sqrt{|\phi|}$ that the latter is equal to (for $\mu = 0$ it vanishes immediately since ϕ does not depend on y^0)

$$\bar{\partial}_i \sqrt{|\phi|} \left(\bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma \sqrt{|\phi|} y^i + 3\delta^i{}_{(\mu} \bar{\partial}_\nu \bar{\partial}_{\sigma)} \sqrt{|\phi|} \right) \tag{10.32}$$

$$= \sqrt{|\phi|} \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma \sqrt{|\phi|} + 3 \bar{\partial}_{(\mu} \sqrt{|\phi|} \bar{\partial}_\nu \bar{\partial}_{\sigma)} \sqrt{|\phi|} \tag{10.33}$$

$$= \tfrac{1}{2} \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma \left(\sqrt{|\phi|} \right)^2 = \tfrac{1}{2} \bar{\partial}_\mu \bar{\partial}_\nu \bar{\partial}_\sigma |\phi| = 0, \tag{10.34}$$

which vanishes because ϕ is quadratic[2]. Hence our modified unicorns are indeed non-Berwaldian Landsberg metrics, justifying their name, and ensuring that the field equation reduces to the (weakly) Landsberg field equation (7.13). The next proposition characterizes Ricci-flat modified unicorn metrics. Below we will see that any vacuum solution to the field equation must be Ricci-flat, hence this result even characterizes all vacuum solutions of modified unicorn type.

Proposition 10.2.1. *F is Ricci-flat if and only if f has the form $f(x^0) = c_1 \exp(c_2 x^0)$, with $c_1, c_2 = const.$*

Proof. By definition, and using homogeneity and the fact that $N^\mu_\nu = \frac{1}{2} \bar{\partial}_\nu G^\mu$, we

[2]Points satisfying $|\phi| = 0$ can be safely ignored since these lie outside the domain $\mathcal{A}$ of the Finsler metric.

have

$$\begin{aligned}
\mathrm{Ric} &= R^{\mu}{}_{\mu\nu}y^{\nu} = (\delta_{\mu}N^{\mu}_{\nu} - \delta_{\nu}N^{\mu}_{\mu})y^{\nu} && (10.35)\\
&= y^{\nu}((\partial_{\mu} - N^{\rho}_{\mu}\bar{\partial}_{\rho})N^{\mu}_{\nu} - (\partial_{\nu} - N^{\rho}_{\nu}\bar{\partial}_{\rho})N^{\mu}_{\mu}) && (10.36)\\
&= y^{\nu}\partial_{\mu}N^{\mu}_{\nu} - y^{\nu}\partial_{\nu}N^{\mu}_{\mu} - y^{\nu}N^{\rho}_{\mu}\bar{\partial}_{\rho}N^{\mu}_{\nu} + y^{\nu}N^{\rho}_{\nu}\bar{\partial}_{\rho}N^{\mu}_{\mu} && (10.37)\\
&= \tfrac{1}{2}\left(y^{\nu}\partial_{\mu}\bar{\partial}_{\nu}G^{\mu} - y^{\nu}\partial_{\nu}\bar{\partial}_{\mu}G^{\mu}\right) && (10.38)\\
&\quad - \tfrac{1}{4}\left(y^{\nu}\bar{\partial}_{\mu}G^{\rho}\bar{\partial}_{\rho}\bar{\partial}_{\nu}G^{\mu} - y^{\nu}\bar{\partial}_{\nu}G^{\rho}\bar{\partial}_{\rho}\bar{\partial}_{\mu}G^{\mu}\right) && (10.39)\\
&= \partial_{\mu}G^{\mu} - \tfrac{1}{2}y^{\nu}\partial_{\nu}\bar{\partial}_{\mu}G^{\mu} - \tfrac{1}{4}\left(\bar{\partial}_{\mu}G^{\rho}\bar{\partial}_{\rho}G^{\mu} - 2G^{\rho}\bar{\partial}_{\rho}\bar{\partial}_{\mu}G^{\mu}\right). && (10.40)
\end{aligned}$$

Using the identities

$$\begin{array}{ll}
\bar{\partial}_0 P = 2f'/f, & \bar{\partial}_0 G^0 = 2y^0 f'/f,\\
\bar{\partial}_0^2 G^0 = 2f'/f, & \bar{\partial}_0\bar{\partial}_i G^0 = \bar{\partial}_0\bar{\partial}_i P = \bar{\partial}_0^2 P = 0,\\
y^i\bar{\partial}_i G^0 = -2|\phi|f'/f, & y^i\bar{\partial}_i P = 2\,\mathrm{sgn}(\phi y^0)\sqrt{|\phi|}f'/f,
\end{array} \qquad (10.41)$$

one finds after some slightly tedious manipulations that the last two terms in the expression for the Ricci tensor can both be expressed as

$$\bar{\partial}_{\mu}G^{\rho}\bar{\partial}_{\rho}G^{\mu} = 2G^{\rho}\bar{\partial}_{\rho}\bar{\partial}_{\mu}G^{\mu} = nP^2, \qquad (10.42)$$

where $n = \dim M = 4$ in our case. Hence these terms cancel each other out precisely. Denoting $G^0 = \bar{G}^0 f'/f$ and $P = \bar{P}f'/f$, so that $\bar{G}^0$ and $\bar{P}$ do not depend on x^{μ}, and using the fact that $\bar{\partial}_{\mu}G^{\mu} = nP$, one finds furthermore that

$$\partial_{\mu}G^{\mu} = (\partial_0^2 \log|f|)\bar{G}^0, \qquad (10.43)$$

$$y^{\nu}\partial_{\nu}\bar{\partial}_{\mu}G^{\mu} = ny^0(\partial_0^2\log|f|)\bar{P}. \qquad (10.44)$$

Consequently, we have

$$\begin{aligned}
\mathrm{Ric} &= \partial_{\mu}G^{\mu} - \tfrac{1}{2}y^{\nu}\partial_{\nu}\bar{\partial}_{\mu}G^{\mu} && (10.45)\\
&= (\partial_0^2\log|f|)\left((1-n)(y^0)^2 - n|y^0|\mathrm{sgn}(\phi)\sqrt{|\phi|} - |\phi|\right), && (10.46)
\end{aligned}$$

which in dimension $n = 4$ reduces to

$$\mathrm{Ric} = -(\partial_0^2\log|f|)\left(3(y^0)^2 + 4|y^0|\mathrm{sgn}(\phi)\sqrt{|\phi|} + |\phi|\right). \qquad (10.47)$$

If $\mathrm{Ric} = 0$ then we must, in particular, have

$$0 = \bar{\partial}_0^2 \mathrm{Ric} = -6(\partial_0^2 \log |f|). \tag{10.48}$$

It thus follows that $\mathrm{Ric} = 0$ if and only if $\partial_0^2 \log |f| = 0$, the general solution to which is given by the stated form of f. □

This shows that the following family of Finsler metrics are exact vacuum solutions to Pfeifer and Wohlfarth's field equation in Finsler gravity:

$$F = \begin{cases} F_0 & \text{if} \quad |y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|} \neq 0 \\ 0 & \text{if} \quad |y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|} = 0 \end{cases}, \tag{10.49a}$$

where

$$F_0 = c_1 e^{c_2 x^0} \left(|y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|} \right) \exp\left(\frac{|y^0|}{|y^0| + \mathrm{sgn}(\phi)\sqrt{|\phi|}} \right). \tag{10.49b}$$

Furthermore, the following proposition shows that *any* solution of the type (10.9) must have this form. In other words, all vacuum solutions of modified unicorn type are in fact Ricci-flat.

Proposition 10.2.2. *A modified unicorn metric* (10.9) *is a solution to the Finslerian field equations in vacuum if and only if* $f(x^0) = c_1 e^{c_2 x^0}$*, i.e. if and only if it can be written as* (10.49)*.*

Proof. In four spacetime dimensions, the proof is straightforward and most easily performed in convenient coordinates in which ϕ is diagonal with all nonvanishing entries equal to $+1$ or -1. Due to the size of some of the expressions involved, however, we only give a sketch of the proof here. From (10.9) one can directly compute $g_{\mu\nu}$ and then its inverse $g^{\mu\nu}$. From (10.47) together with the definition of the Finsler-Ricci tensor (2.34) one can immediately compute $R_{\mu\nu}$. We omit the intermediate expressions because they are somewhat lengthy, but plugging all of this into the (weakly) Landsberg field equation (7.12), we obtain

$$\frac{-\partial_0^2 \log |f|}{3\sqrt{|\phi|}} \left(-4\,\mathrm{sgn}(\phi)|y^0|^3 - 5y^0\sqrt{|\phi|} + |\phi|^{3/2} \right) = 0. \tag{10.50}$$

This equation can only be satisfied for all y^μ for which it is defined if $\partial_0^2 \log |f| = 0$, in which case (10.47) shows that F is in fact Ricci-flat and therefore, by Proposition 10.2.1, must have the form (10.49), as desired. □

10.3 A linearly expanding universe

We now turn to the physical interpretation of one of the unicorn solutions obtained above. We will focus on the case where ϕ has $(-,-,-)$ signature. In that case, we will see that the solution can be interpreted as a linearly expanding (or contracting) universe.

We start by noticing that for any time slice $x^0 = \text{const.}$, $y^0 = 0$, the Finsler metric describing the spatial geometry is a function only of the 3-dimensional Riemannian Finsler Lagrangian $\omega^2 = (y^1)^2 + (y^2)^2 + (y^3)^2$. Since ω^2 is homogeneous and isotropic, the spatial Finsler metric must be as well, for each time slice. This shows that the Finsler spacetime metric has cosmological symmetry, i.e. is spatially homogeneous and isotropic. Alternatively, it has been shown [71, 28] that a Finsler metric has cosmological symmetry if and only if it can be written in the form $F = F(x^0, y^0, \omega)$, where

$$\omega^2 = \frac{(y^r)^2}{1 - kr^2} + r^2 \left((y^\theta)^2 + \sin^2\theta (y^\phi)^2 \right), \qquad k = 0, \pm 1 \tag{10.51}$$

is the standard spatial cosmological line element in coordinates (r, θ, ϕ), represented as a scalar function on TM. With $k = 0$ this corresponds to our ω^2 above, and hence we conclude also in this way that indeed, provided ϕ has $(-,-,-)$ signature, the modified unicorn metric has cosmological symmetry.

Second, it turns out that the modified unicorn metric is conformally flat. More precisely, we can write the Finsler metric

$$F = f(x^0)\bar{F}(y), \tag{10.52}$$

as a conformal factor $f(x^0)$ (which is a function only of the timelike coordinate) times the 'reference' Finsler metric

$$\bar{F} = \left(|y^0| - \sqrt{|\phi(\hat{y})|} \right) e^{\frac{|y^0|}{|y^0| - \sqrt{|\phi(\hat{y})|}}}, \tag{10.53}$$

which is flat. Indeed, since $\bar{F}$ has no x-dependence, it follows from (2.33) that the spray vanishes identically. In other words, the geodesics of $\bar{F}$ are straight lines in these coordinates, and (2.32) implies that the curvature of $\bar{F}$ vanishes, $\bar{N}^\rho_\mu = 0$.

This situation is reminiscent of FLRW cosmology, which also features cosmological symmetry as well as conformal flatness. Indeed, in the simplest (spatially flat) case, the FLRW metric can be written as $\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)$, where $a(t)$ is the scale factor. Introducing the conformal time coordinate $\zeta =$

$\int^t \frac{1}{a(t)} \mathrm{d}t$ the metric takes the form

$$\mathrm{d}s^2 = a(\zeta)^2 \left(-\mathrm{d}\zeta^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right). \tag{10.54}$$

The Finsler metric corresponding to this line-element reads

$$F_{\mathrm{FLRW}}(x,y) = a(\zeta)\sqrt{|-(y^\zeta)^2 + (y^x)^2 + (y^y)^2 + (y^z)^2|} \tag{10.55}$$

$$= a(\zeta)\bar{F}_{\mathrm{FLRW}}(y), \tag{10.56}$$

where $\bar{F}_{\mathrm{FLRW}}$ is the flat 'reference' metric in the FLRW case. Comparing (10.56) to (10.52) suggests that in the Finslerian case, the coordinate x^0 might have the interpretation of a kind of conformal time, while $f(x^0)$ might play a role analogous to the scale factor $a(\zeta)$. Since the transformation $t \to \eta$ is characterized by $\partial t/\partial \eta = a$ and $\partial t/\partial x^i = 0$, it makes sense in this analogy to perform a similar (inverse) coordinate change $x^0 \to \tilde{t}$ according to $\partial \tilde{t}/\partial x^0 = f$ and $\partial \tilde{t}/\partial x^i = 0$ in the Finsler metric. Here $\tilde{t}$ is a new coordinate that, if possible, we would like to interpret as *cosmological time*, analogous to t in the FLRW metric. In fact, we will omit the tilde and just write $\tilde{t} = t$ from here onwards. Then the relation between the coordinates implies that $y^t = fy^0$ and hence the modified unicorn metric attains the form

$$F = \left(|y^t| - f(t)\sqrt{|\phi|}\right) e^{\frac{|y^t|}{|y^t| - f(t)\sqrt{|\phi|}}}. \tag{10.57}$$

where $f(t)$ is shorthand for $f(x^0(t))$. Indeed, it is now clear that:

- $f(t)$ has the interpretation of a scale factor, for the spatial geometry at a time slice $t = t_0$ is given by[3]

  $$F_{\mathrm{sp}} = F\big|_{y^t=0, t=t_0} = f(t_0)\sqrt{(y^1)^2 + (y^2)^2 + (y^3)^2}, \tag{10.58}$$

  which is just the Finsler metric of flat three-dimensional Euclidean space multiplied by $f(t_0)$; equivalently,

  $$\mathrm{d}s^2 = f(t_0)^2 \left((\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2\right). \tag{10.59}$$

- The coordinate t has the interpretation of cosmological time since it cor-

[3] We have removed the irrelevant overall minus sign in order to render the spatial geometry positive definite.

responds (up to a constant[4]) to the proper time of a stationary observer with $\dot{x}^i = 0$. Note that the latter describes a geodesic.

If we additionally require the metric to be a vacuum solution then we know from Proposition 10.2.2 that $f(x^0) = c_1 e^{c_2 x^0}$. In that case it follows from $\partial t/\partial x^0 = f(x^0) = c_1 e^{c_2 x^0}$ that

$$x^0(t) = \frac{1}{c_2} \ln\left(\frac{c_2}{c_1}(t - c_3)\right) , \tag{10.60}$$

where c_3 is another integration constant, and hence we find that the scale factor $f(t)$ as a function of cosmological time t is given by

$$f(t) \equiv f(x^0(t)) = c_2(t - c_3) . \tag{10.61}$$

Thus the modified unicorn vacuum solutions describe a linearly expanding or contracting universe.

As an additional mathematical curiosity, it turns out that these solutions are not only Ricci-flat and conformally flat, as already noted, but flat, in the sense that the nonlinear curvature tensor $R^\rho{}_{\mu\nu} = \delta_\mu N^\rho_\nu - \delta_\nu N^\rho_\mu$ vanishes identically. This follows by explicit computation (which we omit here). Even though $R^\rho{}_{\mu\nu} = 0$, the geometry is nontrivial. This would be impossible in pseudo-Riemannian geometry. Indeed, in pseudo-Riemannian geometry a vanishing curvature tensor would imply that there exist coordinates in which geodesics are straight lines, or, equivalently, in which the spray vanishes identically, $G^\rho = 0$. In the case at hand, though, we are dealing with a Finsler metric that is non-Berwaldian, meaning that, whatever coordinates we use, the spray G^ρ will never be quadratic, and hence, in particular, it will never vanish identically. This proves the nonexistence of a coordinate system in which the geodesics of the unicorn metric are straight lines, showing that the geometry can justly be characterized as nontrivial, even though it has zero curvature.

10.4 Concluding remarks

The results obtained in this chapter motivate a systematic search for cosmological Landsberg spacetimes that solve the field equations. Indeed, while the Berwald

[4] Along the worldline of a stationary observer, which we may parameterize as $x^\mu(t) = (t, x^i_0)$, $\dot{x}^\mu = (1, 0, 0, 0)$, we have $F(x, \dot{x}) = |\dot{x}^t| e^{|\dot{x}^t|/|\dot{x}^t|} = e$, so that $\mathrm{d}\tau/\mathrm{d}t = e$. Hence, strictly speaking, $t = \tau/e$ is a constant multiple of the proper time.

property has been shown to be too restrictive to be able to model nontrivial cosmological dynamics [71], we have shown here that—going up one step in generality—Landsberg spaces do have this capacity. With the recent results characterizing cosmological symmetry in Finsler spacetimes [71] and Elgendi's machinery for constructing unicorns using conformal transformations [123, 49], we have a lot of tools at our disposal to study such cosmological unicorns. This is work in progress.

Discussion and Outlook

The results obtained in this dissertation can be divided into two main (interrelated) categories:

- the characterization of Berwald spaces and their metrizability (Part II);
- vacuum solutions in Finsler gravity (Part III).

Berwald spaces are relevant not only from an intrinsic mathematical point of view but also in the context of Finsler gravity. They represent the first step up in complexity relative to pseudo-Riemannian manifolds, and they are relatively easy to handle, computationally. In Chapter 5 we introduced a new characterization of Berwald spaces and we explored some of its consequences. In particular, we obtained a general necessary and sufficient condition for (α, β)-metrics to be of Berwald type and we applied this Berwald condition to several specific (α, β)-metrics: Randers metrics (for which the result is well known), exponential (α, β)-metrics, m-Kropina metrics (for which the result is also known) and generalized m-Kropina metrics. While based mostly on our published work [H1], we have presented these results here with a substantially heightened focus on mathematical rigor relative to the original article. Apart from the applications discussed here, these results have also found application in the classification of cosmological Berwald spaces [71, 72] and it seems worthwhile to investigate in the future whether there are more possible applications.

While all positive definite Berwald spaces defined on $\mathcal{A} = TM_0$ are pseudo-Riemann metrizable according to Szabo's well-known metrization theorem, it was demonstrated in Chapter 6 that this is not true any longer in the general setting of Berwald spaces of arbitrary signature. We provided a general argument

for this along with a simple explicit counterexample, based on [H2]. Next, we investigated the situation more closely in the case of generalized m-Kropina spaces with closed null 1-form, and (standard) m-Kropina spaces with arbitrary 1-form, and we obtained the precise necessary and sufficient conditions for local metrizability. These results are based in part on, but considerably generalize the findings of [H3].

Interestingly, in all of the cases considered, we concluded that local metrizability is equivalent to the symmetry of the Finsler-Ricci tensor. One of these implications—that a locally metrizable affine connection requires a symmetric Ricci tensor—is trivial, but the opposite implication—that a symmetric Finsler-Ricci tensor implies local metrizability—is quite wonderfully surprising. Of course, this equivalence could be purely incidental and specific to the types of metrics considered here, but one is naturally led to wonder whether such a result might hold more generally. It is thus very much of interest to study the metrizability of other types of Berwald spaces and see whether the equivalence holds up. The ultimate goal would be to prove in complete generality that a Berwald space is locally metrizable if and only if it has a symmetric Finsler-Ricci tensor, or find a counterexample to this conjecture.

Chapter 6 concluded with the local classification of locally metrizable, Ricci-flat m-Kropina spaces in $1 + 3$D. These findings have not been published yet and will appear in a forthcoming article. Two cases can be distinguished: either the 1-form is null or it is not. In the first case, we found that locally, only a single family of such spaces exists, namely with α given by a classical pp-wave metric, and β its defining covariantly constant null 1-form. In the latter case, we found that for each (admissible) value of m there is locally, a unique (necessarily locally metrizable) Ricci-flat m-Kropina space of Berwald type, namely with α given by the flat Minkowski metric and β having constant components in 'flat' coordinates. We have restricted to the scenario that α is of Lorentzian signature, but analogous results can be proven easily in other signatures.

Since Ricci-flatness is a sufficient (and in some cases even necessary) condition for a Berwald space to be a vacuum solution to the Finsler gravity field equations, these results have direct consequences for the classification of vacuum solutions of m-Kropina type. Specifically, apart from the above-mentioned 'trivial' solutions, a Berwald m-Kropina space can only be a vacuum solution if one of the following two conditions is met:

(i) the 1-form is null and F is not locally metrizable;

(ii) the 1-form is not null and F is not Ricci flat.

All that is currently known about such scenarios is that some examples of Ricci-

flat spaces (and hence vacuum solutions) satisfying (i) exist; whether spaces satisfying (ii) exist is unknown. To complete the classification of m-Kropina vacuum solutions of Berwald type, a thorough investigation of both cases (i) and (ii) is required.

And that brings us to Part III, in which vacuum solutions to Pfeifer and Wohlfarth's field equation in Finsler gravity were the central topic. We introduced several classes of exact vacuum solutions, classified them, and discussed their physical interpretation.

To begin with, we proved in Chapter 8 that any (α, β)-metric constructed from a vacuum solution to Einstein's equation α and a covariantly constant 1-form β is a Ricci-flat vacuum solution in Finsler gravity, allowing us to deform solutions in general relativity (GR) to solutions in Finsler gravity. We then classified such solutions and showed that all nontrivial ones are generalizations of the well-known general relativistic pp-wave solutions. For Randers metrics as well as modified Randers metrics—a slightly different version of the Randers metric that we introduced and studied in depth because of its preferable causal properties—we even demonstrated that all vacuum solutions of Berwald type must be of this form, classifying (locally) all Berwald-Randers vacuum solutions. These results are based on [H4] and [H5].

While such classification results are intrinsically valuable from a mathematical point of view, it is essential from a physics perspective to understand what kind of physics such solutions represent. That is why, in Chapter 9, we investigated the observational signature of our pp-wave type Finsler spacetimes, based on [H5]. After a physically motivated twofold linearization procedure—linearizing both in the departure from flatness as well as in the departure from GR—we found that these solutions can be interpreted as Finslerian gravitational waves and we were able to obtain an explicit expression for the radar distance, the main observable in gravitational wave interferometry. The result was quite surprising: if such a Finslerian gravitational wave were to pass the Earth, its effect on an interferometer would be indistinguishable from a similar general relativistic gravitational wave!

This is on the one hand somewhat disappointing, as it suggests that interferometer measurements may not have the power to reveal possible Finslerian characteristics of our universe. On the other hand, it also means that current gravitational wave measurements are all compatible with the idea that spacetime has a Finslerian nature. Moreover, while research in Finsler gravity is generally quite abstract, our radar distance result—which may be viewed not in the least as a proof of concept—shows how clear physical predictions can be extracted

from an *a priori* abstract Finslerian model of spacetime. Repeating such an analysis in other scenarios (e.g. more generic types of Finslerian gravitational waves, higher-order analyses, or completely different Finsler spacetimes) may lead to additional insight into the observational signature of Finsler gravity.

Another highly interesting direction for future research in this context would be to investigate whether Finslerian effects also play a role in the generation of gravitational waves during, say, a black hole merger event. It is to be expected that such effects would lead to a Finslerian correction to the waveform and this *could*, in principle, be measured in interferometer experiments, especially since our results show how the waveform correlates with the experimental results—namely in exactly the same way in which the waveform of a general relativistic gravitational wave does. To investigate this, Finslerian black hole solutions first need to be better understood, and research in this direction is actively ongoing, in part by the author. All 4-dimensional spherically symmetric Finsler metrics of Berwald type have recently been classified [121] and, while a Birkhoff-type theorem has been proven for Ricci-flat spherically symmetric Berwald spacetimes [102], it will be demonstrated in a forthcoming article that non–Ricci-flat spherically symmetric vacuum solutions more general than the Schwarzschild geometry do exist, suggesting that the gravitational field around a black hole might indeed feature Finslerian effects.

While all solutions discussed up to this point have been of Berwald type, we introduced and studied a (non-Berwaldian Landsberg) unicorn solution to the vacuum field equation in Chapter 10, which is based on [H6]. Its mathematical interest lies in the fact that this is the first and only non-Berwaldian solution to Pfeifer and Wohlfarth's equation currently known in the literature and that it is of unicorn type, but in addition to that, we have seen that it can be interpreted very naturally in a cosmological setting as well. Indeed, we have shown that the solution has cosmological symmetry, i.e. is spatially homogeneous and isotropic, that it has a well-behaved light cone structure, and that its (flat) spatial slices evolve in cosmological time by means of a scale factor—reminiscent of the Friedmann-Lemaître-Robertson-Walker (FLRW) metric—which has a linear dependence on cosmological time. This leads to the natural interpretation of our unicorn solution as describing a linearly expanding (or contracting) Finslerian universe.

As an additional curiosity, we found that the requirement of a physically viable light cone structure does not, strictly speaking, necessitate Lorentzian signature, as is often assumed. This is illustrated by the signature anomaly of our above-mentioned cosmological unicorn solution, which indeed has positive definite signature everywhere on its domain, and yet has a light cone structure

that is equivalent to that of flat Minkowski space. This surprising observation raises the question of how strictly one should adhere to the standard requirement of Lorentzian signature in the context of Finsler spacetime geometry. Of course, there are other reasons why Lorentzian signature might still be a desirable property—extremization properties of geodesics, existence, and uniqueness (in the right context) of solutions to the field equations, the propagation speed of the gravitational interaction, and more—but nevertheless it is worth thinking about this question in depth.

Our results also motivate a systematic search for cosmological Landsberg spacetimes that solve the field equations. Since (properly Finslerian) cosmological solutions of Berwald type are necessarily static [71] and hence not particularly interesting, any nontrivial such Landsberg spacetime must necessarily be a unicorn. With the recent results characterizing cosmological symmetry in Finsler spacetimes [71], Elgendi's machinery for constructing unicorns using conformal transformations [123, 49], and the discovery of the first unicorn solution discussed here, we are very optimistic about this endeavor. This is work in progress.

Another important next step in the study of Finsler gravity, not only in the context of cosmology but in general, is the investigation of solutions to the field equation in the presence of matter. Promising ideas about the Finslerian description of matter have arisen recently. For instance, taking into account the individual motion of each gas particle in a kinetic gas, one obtains an energy-momentum tensor that is direction-dependent [88, 89]. Such an energy-momentum tensor cannot be accommodated in GR for obvious reasons, and indeed, in the usual Einstein-Vlasov treatment [90, 91] the individual motion of the gas particles is averaged over. Finsler gravity, however, is perfectly well capable of incorporating this direction-dependence and hence it could very well be that Finsler gravity is able to provide a more accurate description of the gravitational field of a kinetic gas. Currently, matter solutions in Finsler gravity are still completely unexplored, but with the before-mentioned ideas in mind, it seems a very promising avenue.

Finally, let us zoom out and return to the broader perspective. Fundamental physics stands before an enormous challenge: the unification of gravity with quantum mechanics. Ultimately, the research covered in this dissertation stands in service of this goal. Indeed, the possibility, suggested by quantum gravity research, that local Lorentz invariance—which lies at the heart of GR—may not be fundamental leads directly to the notion of Finsler spacetime geometry, since the latter beautifully accommodates departures from Lorentz invariance already at the classical level. This has been the primary motivation for our

work. Of course, the topics covered here represent only a tiny fraction of the larger picture. Yet, we are hopeful that our results, though modest in scope, will prove instrumental in the understanding of gravity and that, eventually, they may even shine light on the reconciliation of gravity with quantum mechanics. With this grand prospect, we have come to the end of this dissertation.

Appendices

APPENDIX A

Special Pseudo-Riemannian Metrics

In this appendix we discuss several special types of pseudo-Riemannian metrics and their standard expression in coordinates, starting with metrics that admit a closed null 1-form. But first, a remark about the notation is in order. In several places below, b will *a priori* denote a 1-form. But we will occasionally use the same symbol b to denote its vector equivalent via the metric-induced isomorphism. For example, $b = \mathrm{d}u$ means that $b_i = \partial_i u$ and $b = \partial_v$ means that $b^i = \partial x^i/\partial v$. No confusion should be possible as to what is meant.

A.1 Metrics admitting a closed null 1-form

Proposition A.1.1. *Any pseudo-Riemannian metric admitting a closed null 1-form b is locally (whenever $b \neq 0$) given by*

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_b\mathrm{d}u\mathrm{d}x^b + h_{bc}\mathrm{d}x^b\mathrm{d}x^c, \tag{A.1}$$

in coordinates (u, v, x^a), $a = 3, \ldots, n$, with $b = \mathrm{d}u$, where H, W_a are smooth functions, and where, for fixed u and v, h_{ab} is a pseudo-Riemannian metric of dimension $n - 2$.

Proof. First, we may pick coordinates $(v, x^2, \ldots, x^n)$ around p adapted to b in the sense that $b = \partial_v$, i.e. $b^i = \delta^i_1$. At this point, the metric has the general form $a = a_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j$. The null character of b manifests as the fact that $a_{11} =$

$a_{vv} = 0$ in these coordinates. Because b is closed and hence locally exact, we may write, locally, $b = \mathrm{d}u$ for some function $u(v, x^2, \dots, x^n)$. Equivalently, $b_i = \partial_i u$. Note also that $\partial_i u = b_i = a_{ij} b^j = a_{ij}\delta^j_1 = a_{i1}$. Since $a_{11} = 0$, it follows that $\partial_v u = \partial_1 u = 0$. As $b \neq 0$ by assumption, there must be some $i \geq 2$ such that $\partial_i u = a_{i1} \neq 0$ in a neighborhood of p. Order the coordinates $x^2, \dots, x^n$ such that this is true for $i = 2$, i.e. assume without loss of generality that $a_{21} \neq 0$. Next, define the map

$$x = (v, x^2, \dots, x^n) \mapsto \tilde{x} = (v, u(x^2, \dots, x^n), x^3, \dots, x^n). \tag{A.2}$$

Its Jacobian matrix and its inverse are given by

$$J^i{}_j = \frac{\partial \tilde{x}^i}{\partial x^j} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{21} & a_{31} & \dots & a_{n1} \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{A.3}$$

$$(J^{-1})^i{}_j = \frac{\partial x^i}{\partial \tilde{x}^j} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1/a_{21} & -a_{31}/a_{21} & \dots & a_{n1}/a_{21} \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{A.4}$$

and since $\det J = a_{21} \neq 0$ this matrix is invertible, so $x \mapsto \tilde{x}$ is a local diffeomorphism at p. It remains to find the form of the metric in the new coordinates. We have

$$\tilde{a}_{ij} = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial x^\ell}{\partial \tilde{x}^j} a_{k\ell}, \qquad \text{i.e.} \qquad \tilde{a} = J^{-1T} a J^{-1}. \tag{A.5}$$

Therefore we have

$$\tilde{a}_{11} = (J^{-1T})_1{}^i a_{ij} (J^{-1})^j{}_1 = a_{11} = 0, \tag{A.6}$$

$$\tilde{a}_{12} = (J^{-1T})_1{}^i a_{ij} (J^{-1})^j{}_2 = 1, \tag{A.7}$$

$$\tilde{a}_{1b} = (J^{-1T})_1{}^i a_{ij} (J^{-1})^j{}_b = a_{12}(-a_{b1}/a_{21}) + a_{1b} = 0, \quad b = 3, \dots, n. \tag{A.8}$$

This shows that $a = \tilde{a}_{ij} \mathrm{d}\tilde{x}^i \mathrm{d}\tilde{x}^j = 2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_b \mathrm{d}u\mathrm{d}x^b + h_{bc}\mathrm{d}x^b\mathrm{d}x^c$ for certain functions H, W_a, h_{ab}, and hence after a redefinition $v \to -v$ we may

write the metric in the form

$$a = -2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_b\mathrm{d}u\mathrm{d}x^b + h_{bc}\mathrm{d}x^b\mathrm{d}x^c. \tag{A.9}$$

It follows from the easily checked fact that $\det h = -\det a \neq 0$ that h_{ab} is itself a pseudo-Riemannian metric of dimension $n-2$. □

In matrix form, such a metric reads, schematically,

$$g = \begin{pmatrix} H & -1 & W_a \\ -1 & 0 & 0 \\ W_a & 0 & h_{ab} \end{pmatrix} \tag{A.10}$$

with inverse is given by

$$g^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & -H + h^{ab}W_aW_b & h^{ab}W_b \\ 0 & h^{ab}W_b & h^{ab} \end{pmatrix} \tag{A.11}$$

$$= \begin{pmatrix} 0 & -1 & 0 \\ -1 & -H + W^2 & W^a \\ 0 & W^a & h^{ab} \end{pmatrix}, \tag{A.12}$$

where h^{ab} is the inverse of h_{ab}, indices have been raised with h, and $W^2 = h^{ab}W_aW_b$. We note that the Christoffel symbols with upper index u have a particularly simple form,

$$\Gamma^u_{ij} = \tfrac{1}{2}\partial_v g_{ij}. \tag{A.13}$$

A.2 Metrics admitting a parallel vector field

Next, we discuss pseudo-Riemannian metrics that admit a parallel (i.e. covariantly constant) vector field. Note that the existence of a parallel vector field is equivalent to the existence of a parallel 1-form of the same causal character. We distinguish the two cases that the 1-form is null, or that the 1-form is not null, starting with the former. In this case, the resulting metric is called a CCNV[1] metric or a pp-wave metric.

[1] CCNV stands for *covariantly constant null vector*.

Proposition A.2.1. *Any pseudo-Riemannian metric admitting a parallel null 1-form b_i is locally (whenever $b \neq 0$) given by*

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H(u, x^a)\mathrm{d}u^2 + 2W_b(u, x^a)\mathrm{d}u\mathrm{d}x^b + h_{bc}(u, x^a)\mathrm{d}x^b\mathrm{d}x^c, \quad \text{(A.14)}$$

in coordinates (u, v, x^a), $a = 3, \ldots, n$, with $b = \mathrm{d}u$, and where, for each fixed u, h_{bc} is a pseudo-Riemannian metric of dimension $n-2$. Conversely, $\mathrm{d}u$ is null and parallel with respect to any metric of the form (A.14).

Proof. It is easy to see that a 1-form is parallel if and only if it is closed and satisfies the Killing equation. Hence b_i is in particular closed and it follows from Proposition A.1.1 that locally we may write

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H\mathrm{d}u^2 + 2W_b\mathrm{d}u\mathrm{d}x^b + h_{bc}\mathrm{d}x^b\mathrm{d}x^c, \qquad b = \mathrm{d}u. \quad \text{(A.15)}$$

Computing $\nabla_i b_j$ explicitly in these coordinates, using the fact that $b_i = \delta_i^u$ and $g^{ui} = -\delta_v^i$ and $g_{iv} = 0$, yields

$$\nabla_i b_j = -\frac{1}{2}\frac{\partial a_{ij}}{\partial v}. \quad \text{(A.16)}$$

Since b_i is parallel, all these components must vanish, and consequently all metric functions must be independent of v. The converse statement that $\mathrm{d}u$ is always null and parallel is straightforward to check explicitly. □

If, on the other hand, the 1-form is not null, the metric decomposes locally as a direct product of a flat 1-dimensional metric and an arbitrary $(n-1)$-dimensional metric. In order to prove this we employ the Frobenius theorem.

Recall that a distribution is called involutive if it is closed under the Lie bracket, and it is called completely integrable (we follow the terminology of e.g. [55]) if it is spanned by a set of coordinate vector fields in some cubic chart. The Frobenius theorem states that any involutive distribution is completely integrable. We will take this result for granted.

The following is a consequence of a more general result (see e.g. Appendix C of [124]), but here we give a self-contained proof directly from the Frobenius theorem. To ease the notation, coordinate labels will momentarily run from 0 to $n-1$ rather than 1 to n, even though the signature is not necessarily Lorentzian but in principle arbitrary. Greek indices $\mu, \nu, \rho, \ldots$ run from 0 to $n-1$, whereas Latin indices $i, j, k, \ldots$ run from 1 to $n-1$.

Proposition A.2.2. *Any pseudo-Riemannian metric admitting a nonnull (i.e. $|b|^2 \neq 0$) parallel vector field b^μ can locally be written as a product metric*

$$\mathrm{d}s^2 = \pm(\mathrm{d}x^0)^2 + h_{ij}(x^k)\mathrm{d}x^i\mathrm{d}x^j. \tag{A.17}$$

where $b = b_0\,\mathrm{d}x^0$ with $b_0 =$ constant, and where h_{ij} is an $(n-1)$-dimensional pseudo-Riemannian metric.

Proof. We apply the Frobenius theorem twice. First to the distribution spanned by b^μ, and then to the distribution orthogonal to b^μ. Since any vector commutes with itself, it follows immediately that $\mathrm{span}\{b\}$ is involutive and hence completely integrable, so there exist coordinates y^μ such that $b = \partial/\partial y^0$. We keep this coordinate system in mind.

Next, consider the distribution D consisting of all vectors orthogonal to b^μ. Note that since $|b|^2 \neq 0$, by assumption, the orthogonal complement $b^\perp$ of b in T_xM is an $(n-1)$-dimensional linear subspace, and symbolically $T_xM = \mathrm{span}\{b\} \oplus b^\perp$. Hence D is an $(n-1)$-dimensional distribution. Now let $v, w \in D$. Then, by torsion-freeness, we have $[v,w] = \nabla_v w - \nabla_w v$, and by metric-compatibility, $\nabla g = 0$, we have $g(\nabla_v w, b) = -g(w, \nabla_v b)$ and $g(\nabla_w v, b) = g(v, \nabla_w b)$. Hence, since b is parallel we see that

$$\begin{aligned} g([v,w],b) &= g(\nabla_v w, b) - g(\nabla_w v, b) && \text{(A.18)}\\ &= -g(w, \nabla_v b) + g(v, \nabla_w b) && \text{(A.19)}\\ &= 0. && \text{(A.20)} \end{aligned}$$

This shows that $[v,w] \in D$ and hence D is involutive. By the Forbenius theorem, the $(n-1)$-dimensional distribution D is therefore completely integrable, that is, there exist local coordinates $z^0, \ldots, z^{n-1}$ such that D is spanned by $\partial/\partial z^1, \ldots, \partial/\partial z^{n-1}$.

Now we combine the two coordinate systems y^μ and z^μ. Namely, we define new coordinates x^μ by setting $x^0 = z^0$ and $x^i = y^i$ for $i = 1, \ldots, n-1$. This is a valid coordinate transformation as long as the $\mathrm{d}x^\mu$ are linearly independent. To see that the latter is the case, suppose that at some point in M we have $c_\mu \mathrm{d}x^\mu = 0$, i.e.

$$c_0 \mathrm{d}z^0 = -c_i \mathrm{d}y^i. \tag{A.21}$$

Then, by construction, we see that the LHS annihilates $D = b^\perp$, whereas the RHS annihilates b. Hence the vector corresponding to this 1-form would be contained

in both $\mathrm{span}(b)$ and $b^\perp$. But, since $|b|^2 \neq 0$, we have $\mathrm{span}(b) \cap b^\perp = \{0\}$, so it follows that we must have $c_0 = c_i = 0$. Hence x^μ are valid coordinates on M.

By construction, b is a multiple of $\partial/\partial x^0$. This is because $b = \partial/\partial y^0$, meaning that $\mathrm{d}y^i(b) = 0$ for each i. Therefore $\mathrm{d}x^i(b) = \mathrm{d}y^i(b) = 0$.

Next, we will prove that in the x^μ coordinates, $g_{0i} = 0$. First notice that if a 1-form ω annihilates the orthogonal complement of b, i.e. if $\omega(v) = 0$ for all v satisfying $g_{\mu\nu}v^\mu b^\nu = 0$, then we must have $\omega = \omega_0 \mathrm{d}z^0$, which is clear from the form of D in z^μ-coordinates. But we have $\mathrm{d}x^0 = \mathrm{d}z^0$ and hence it follows as well that if a 1-form ω annihilates the orthogonal complement of b then $\omega = \omega_0 \mathrm{d}x^0$, or equivalently $\omega_i = 0$. We know, however, that $\omega_\mu = g_{\mu\nu}b^\nu$ trivially annihilates the orthogonal complement of b. This therefore implies that $\omega_i = g_{i\nu}b^\nu = 0$. Now recall that $b^\mu = b^0\partial_0$. It follows that $\omega_i = g_{i0} = 0$, as claimed.

Since the cross terms in the metric vanish and b is a multiple of $\partial/\partial x^0$, we also have $b = b_0 \mathrm{d}x^0$. It then follows from the fact that b is closed that $b_0 = b_0(x^0)$.

So far what we have shown is that the metric and 1-form have the form

$$\mathrm{d}s^2 = g_{00}(x^\mu)(\mathrm{d}x^0)^2 + g_{ij}(x^\mu)\mathrm{d}x^i\mathrm{d}x^j, \qquad b = b_0(x^0)\mathrm{d}x^0 \tag{A.22}$$

Next, we explicitly evaluate the equation $\nabla_\mu b_\nu = 0$ in the new coordinates. The ij-component shows that $\nabla_i b_j = -b_0\Gamma^0_{ij} = b_0 g^{00}\partial_0 g_{ij} = 0$, hence $\partial_0 g_{ij} = 0$. And the $0i$-component shows that $\nabla_0 b_i = -b_0\Gamma^0_{0i} = -b_0 g^{00}\partial_i g_{00} = 0$, hence $\partial_i g_{00} = 0$. From the latter it follows that we can do a final coordinate transformation $x^0 \mapsto \bar{x}^0$ such that $\mathrm{d}\bar{x}^0 = \sqrt{|g_{00}(x^0)|}\mathrm{d}x^0$. Then, after dropping the bar over x^0, the metric attains the form

$$\mathrm{d}s^2 = \pm(\mathrm{d}x^0)^2 + g_{ij}(x^k)\mathrm{d}x^i\mathrm{d}x^j. \tag{A.23}$$

Then $\Gamma^0_{00} = 0$ and hence $\nabla_0 b_0 = \partial_0 b_0 = 0$ implies that $b_0 = \text{const}$. Finally, since $\det g_{\mu\nu} = \pm \det g_{ij}$, the $(n-1)$-dimensional metric g_{ij} is nondegenerate and hence pseudo-Riemannian. This completes the proof. □

Corollary A.2.3. *In 4 spacetime dimensions, the only vacuum solution to Einstein's field equations admitting a nonnull parallel vector field or 1-form is (locally) flat Minkowski spacetime.*

Proof. Let $g_{\mu\nu}$ be such a metric. By Proposition A.2.2, it can be written locally as

$$\mathrm{d}s^2 = \pm(\mathrm{d}x^0)^2 + h_{ij}(x^k)\mathrm{d}x^i\mathrm{d}x^j. \tag{A.24}$$

It is a simple exercise to show that the Ricci curvature decomposes as

$$R_{00} = R_{0i} = 0, \qquad R_{ij} = \tilde{R}_{ij}, \tag{A.25}$$

where $\tilde{R}_{ij}$ is the Ricci tensor of h_{ij}. If $g_{\mu\nu}$ is a vacuum to Einstein's field equations, $R_{\mu\nu} = 0$, it therefore follows that $\tilde{R}_{ij} = 0$ and h_{ij} is Ricci-flat. However, in three dimensions it is a standard result that any Ricci-flat pseudo-Riemannian metric is flat, which follows from the Ricci-decomposition of the Riemann tensor because the Weyl tensor vanishes identically. Hence h_{ij} is flat. As the direct product of two flat pseudo-Riemannian manifolds, $g_{\mu\nu}$ is itself flat. □

See also [79, §35.1.1] for an alternative argument and related results.

A.3 Some properties of the pp-wave metric

Consider the pp-wave metric (A.14) in 4 spacetime dimensions,

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H(u, x^a)\mathrm{d}u^2 + 2W_b(u, x^a)\mathrm{d}u\mathrm{d}x^b + h_{bc}(u, x^a)\mathrm{d}x^b\mathrm{d}x^c, \tag{A.26}$$

$a, b, c, \dots = 1, 2$. Under the assumption that this is a vacuum solution to Einstein's field equations, i.e. that the metric is Ricci-flat, it follows from the results in [86] (see also section 24.5 in [79]) that the functions W_a can be transformed away and h_{ab} can be transformed into δ_{ab} by means of an appropriate coordinate transformation[2]. Such a vacuum pp-wave metric can thus be written in standard form

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H(u, x, y)\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2. \tag{A.27}$$

Then the only nonvanishing component of the Ricci tensor is $R_{uu} = -\frac{1}{2}\Delta H$, where $\Delta = \partial_x^2 + \partial_y^2$, which yields the following.

Proposition A.3.1. *A pseudo-Riemannian metric in* $(1{+}3)D$ *admitting a parallel null 1-form is Ricci-flat, i.e. is a vacuum solution to Einstein's field equation, if and only if it can be written locally as*

$$\mathrm{d}s^2 = -2\mathrm{d}u\mathrm{d}v + H(u, x, y)\mathrm{d}u^2 + \mathrm{d}x^2 + \mathrm{d}y^2, \qquad \Delta H = 0. \tag{A.28}$$

[2] In [86], or rather its English translation [125], the term *normal hyperbolic* is used, which is meant to indicate that the metric is *of Lorentzian signature*.

APPENDIX B

A Basic Linearity Result

The aim of this appendix is to prove a basic result, Lemma B.0.2, concerning linearity and constancy of certain functions on $\mathbb{R}^n$, used in Theorem 5.3.12 and Theorem D.2.2. In what follows, f and g will be real-valued functions an open subset of $\mathbb{R}^n$. Moreover, to prevent confusion, we stress that indices $i, j, k, \ldots$ in expressions like x^i will always be indices, not powers, i.e. x^2 denotes the second component of the vector $x \in \mathbb{R}^n$.

Lemma B.0.1. *If the function $x \mapsto f(x)x^i$ is affine for each i, i.e. $f(x)x^i = c^i_k x^k + d^i$, on an open subset $U \subset \mathbb{R}^n$ with $n \geq 2$ such that $0 \notin U$, then $f(x)$ is constant on U.*

Proof. Multiplying $f(x)x^i = c^i_k x^k + d^i$ with x^j and $f(x)x^j = c^j_k x^k + d^j$ with x^i shows that for all i, j we must have

$$c^i_k x^k x^j + d^i x^j = c^j_k x^k x^i + d^j x^i, \tag{B.1}$$

which implies that

$$d^i x^j = d^j x^i, \qquad c^i_k x^k x^j = c^j_k x^k x^i. \tag{B.2}$$

The first equation implies (e.g. by differentiation) that $d^i = 0$ for each i. The second equation can be written as

$$c^i_k \delta^j_l x^k x^l = c^j_k \delta^i_l x^k x^l, \tag{B.3}$$

and hence, by differentiating twice, we obtain that

$$c^i_k\delta^j_l + c^i_l\delta^j_k = c^j_k\delta^i_l + c^j_l\delta^i_k. \tag{B.4}$$

Taking any $i \neq j = k = l$ (this is where the assumption $n \geq 2$ is needed) this reduces to $2c^i_k = 0$. Hence the matrix of coefficients c^i_j must be diagonal. Next, taking any $i = k \neq l = j$, the equation reduces to $c^i_i = c^j_j$ (no summation). Hence $c^1_1 = \cdots = c^n_n =: c$. Substituting this in the original assumption that $f(x)x^i = c^i_k x^k + d^i$ yields $f(x)x^i = cx^i$ and hence, since $x \neq 0$ on U, this implies that $f(x) = c$ is constant. □

Lemma B.0.2. *If g is differentiable on an open subset $U \subset \mathbb{R}^n$ with $n \geq 3$ and if the function $x \mapsto f(x)b^i + g(x)x^i$ is linear for each i, i.e. $f(x)b^i + g(x)x^i = c^i_j x^j$, then $f(x)$ is linear and $g(x)$ is constant.*

Proof. Consider b^i and x^i as the components of two vectors $\vec{b}$ and $\vec{x}$ in $\mathbb{R}^n$ and pick a basis such that $\vec{b} = (1, 0 \ldots, 0)$ in that basis. The assumption that $f(x)b^i + g(x)x^i$ is linear then splits into two, namely:

(a) $f(x) + g(x)x^1$ is linear;

(b) $g(x)x^i$ is linear for all $i > 1$.

Writing $x = (x^1, \bar{x})$, condition (b) implies that, for any x^1 and any $i = 1, \ldots, n-1$, the map

$$\mathbb{R}^{n-1} \ni \bar{x} \mapsto g(x^1, \bar{x})\bar{x}^i = g(x)x^{i+1} = c^{i+1}_j x^j = c^{i+1}_1 x^1 + \sum_{j=1}^{n-1} c^{i+1}_{j+1}\bar{x}^j \tag{B.5}$$

is affine on the open set $U_{x^1} = \{\bar{x} \in \mathbb{R}^{n-1} : (x^1, \bar{x}_0) \in U\}$. Hence, by Lemma B.0.1, this map is constant on $U_{x^1} \setminus \{0\}$, but this implies, by continuity, that the map is constant on U_{x^1}. This means that $g(x)$ can only depend on x^1, i.e. we may write $g(x) = h(x^1)$. Now consider the specific case of $i = 2$. According to (b) and the result just obtained, $h(x^1)x^2$ is linear, that is, we can write $h(x^1)x^2 = c^2_j x^j$. Differentiation with respect to x^2 then shows that $h(x^1) = c^2_2 =$ constant, and hence $g(x) =$ constant. Plugging this into (a) immediately implies that $f(x)$ is linear, completing the proof. □

APPENDIX C

Determinant of an Indefinite (α, β)-Metric

Here we derive the formula (see (C.14) below) for the determinant of a not necessarily positive definite (α, β)-metric. This formula, which appeared originally in [H5], is a generalization of a well-known result for positive definite (α, β)-metrics, and the proof is not substantially different from the positive definite analog. We consider Finsler metrics of the form $F = \alpha\phi(s)$, where $s = \beta/\alpha$, $\alpha = \sqrt{|A|} = \sqrt{|a_{ij}y^i y^j|}$, $A = a_{ij}y^i y^j = \mathrm{sgn}(A)\alpha^2$, and where a_{ij} is assumed to be a pseudo-Riemannian metric, i.e. not necessarily Riemannian/positive definite.

In complete analogy with the positive definite case, it can be shown by direct calculation that the fundamental tensor $g_{ij} \equiv \frac{1}{2}\bar{\partial}_i\bar{\partial}_j F^2$ is given by

$$g_{ij} = \mathrm{sgn}(A)\rho a_{ij} + \rho_0 b_i b_j + \rho_1(b_i\alpha_j + \alpha_i b_j) + \rho_2\alpha_i\alpha_j, \tag{C.1}$$

where we have defined $\alpha_i = a_{ij}y^j/\alpha$, and with coefficients given by

$$\rho = \phi(\phi - s\phi'), \tag{C.2}$$
$$\rho_0 = \phi\phi'' + \phi'\phi', \tag{C.3}$$
$$\rho_1 = -(s\rho_0 - \phi\phi') = -\left[s(\phi\phi'' + \phi'\phi') - \phi\phi'\right], \tag{C.4}$$
$$\rho_2 = -s\rho_1 = s\left[s(\phi\phi'' + \phi'\phi') - \phi\phi'\right]. \tag{C.5}$$

The only difference here with the positive definite case is the factor $\mathrm{sign}(A)$ appearing in the first term in (C.1). Denoting $\dim M = n$, we can write this in

matrix notation as

$$g = \mathrm{sgn}(A)\rho\left(a + UWV^T\right), \tag{C.6}$$

in terms of the three matrices

$$W = \frac{\mathrm{sgn}(A)}{\rho}\mathbb{I}_{4\times 4}, \qquad U = (\vec{b}, \vec{b}, \vec{\alpha}, \vec{\alpha}), \qquad V = (\rho_0\vec{b}, \rho_1\vec{\alpha}, \rho_1\vec{b}, \rho_2\vec{\alpha}). \tag{C.7}$$

Note that U and V are both $n \times 4$ matrices and W is a 4×4 matrix. It is a well-known result (one of the matrix determinant lemmas, see e.g. [117]) that if a is an invertible matrix (which it is in our case) then the determinant of the expression in brackets in (C.6) is equal to

$$\det\left(a + UWV^T\right) = \det\left(\mathbb{I}_{4\times 4} + WV^T a^{-1}U\right)\det a. \tag{C.8}$$

It follows that

$$\det g = \mathrm{sgn}(A)^n\rho^n \det\left(\mathbb{I}_{4\times 4} + WV^T a^{-1}U\right)\det a. \tag{C.9}$$

The matrix product $WV^T a^{-1}U = \frac{\mathrm{sgn}(A)}{\rho}V^T a^{-1}U$ can be evaluated by explicit computation and reads

$$WV^T a^{-1}U = \frac{\epsilon}{\rho}\begin{pmatrix} |b|^2\rho_0 & |b|^2\rho_0 & \epsilon s\rho_0 & \epsilon s\rho_0 \\ \epsilon s\rho & \epsilon s\rho & \epsilon\rho & \epsilon\rho \\ |b|^2\rho_1 & |b|^2\rho_1 & \epsilon s\rho_1 & \epsilon s\rho_1 \\ \epsilon s\rho_2 & \epsilon s\rho_2 & \epsilon\rho_2 & \epsilon\rho_2 \end{pmatrix}, \tag{C.10}$$

where we have written $\epsilon = \mathrm{sgn}(A)$. Hence we obtain

$$\det g = \epsilon^n\rho^n \det a \det\left(\mathbb{I}_{4\times 4} + \frac{\epsilon}{\rho}\begin{pmatrix} |b|^2\rho_0 & |b|^2\rho_0 & \epsilon s\rho_0 & \epsilon s\rho_0 \\ \epsilon s\rho & \epsilon s\rho & \epsilon\rho & \epsilon\rho \\ |b|^2\rho_1 & |b|^2\rho_1 & \epsilon s\rho_1 & \epsilon s\rho_1 \\ \epsilon s\rho_2 & \epsilon s\rho_2 & \epsilon\rho_2 & \epsilon\rho_2 \end{pmatrix}\right) \tag{C.11}$$

$$= \epsilon^n\rho^n \det a \det\left(\mathbb{I}_{4\times 4} + \frac{1}{\rho}\begin{pmatrix} \epsilon|b|^2\rho_0 & \epsilon|b|^2\rho_0 & s\rho_0 & s\rho_0 \\ s\rho & s\rho & \rho & \rho \\ \epsilon|b|^2\rho_1 & \epsilon|b|^2\rho_1 & s\rho_1 & s\rho_1 \\ s\rho_2 & s\rho_2 & \rho_2 & \rho_2 \end{pmatrix}\right) \tag{C.12}$$

$$= \phi^{n+1}(\phi - s\phi')^{n-2}(\phi - s\phi' + (\epsilon|b|^2 - s^2)\phi'')\det a_{ij}. \tag{C.13}$$

Some useful identities that we have used are: $\alpha_i = \mathrm{sgn}(A)y_i/\alpha$ so that $\alpha_i\alpha^i =$

$\mathrm{sgn}(A)$ and $\alpha_i b^i = \mathrm{sgn}(A)s$. We conclude that

$$\boxed{\det g = \phi^{n+1}(\phi - s\phi')^{n-2}(\phi - s\phi' + (\mathrm{sgn}(A)|b|^2 - s^2)\phi'') \det a.} \tag{C.14}$$

In the case that α is positive definite, $\mathrm{sign}(A) = 1$ everywhere, so the formula reduces to the standard result (see e.g. [110]).

APPENDIX D

Additional Proofs

D.1 Metric-compatibility of a nonlinear connection

Proposition D.1.1. *Let (M, L) be a Finsler space and N a torsion-free, homogeneous nonlinear connection on M. TFAE:*

(i) $\delta_i L = 0$,

(ii) $\boldsymbol{\nabla}_D g_{ij} = 0$.

Proof. The implication $(i) \Rightarrow (ii)$ can be checked explicitly using the expression (2.23) for the canonical connection. In terms of the spray

$$G^k = g^{kj} G_j, \qquad G_j = \tfrac{1}{2}\left(y^m \partial_m \bar{\partial}_j L - \partial_j L\right), \tag{D.1}$$

we simply have $N_i^k = \frac{1}{2}\bar{\partial}_i G^k$, by the fundamental lemma of Finsler geometry. We need to show that

$$\boldsymbol{\nabla}_D g_{ij} = y^k \delta_k g_{ij} - N_i^k g_{kj} - N_j^k g_{ik} = 0. \tag{D.2}$$

To that end, consider the fact that we can write

$$N_i^k g_{kj} = \tfrac{1}{2}\bar{\partial}_i \left(g^{\ell k} G_\ell\right) g_{kj} = \tfrac{1}{2}\left(\bar{\partial}_i G_j + g_{kj}\bar{\partial}_i g^{\ell k} G_\ell\right) = \tfrac{1}{2}\left(\bar{\partial}_i G_j - \bar{\partial}_i g_{kj} G^k\right). \tag{D.3}$$

By noting that the two individual terms can be written as

$$\bar{\partial}_i G_j = \partial_{[i}\bar{\partial}_{j]}L + y^m \partial_m g_{ij}, \qquad G^k = N^k_m y^m, \tag{D.4}$$

where the second one follows from Euler's theorem, it follows that

$$\begin{aligned} N^k_i g_{kj} &= \tfrac{1}{2}\left(\partial_{[i}\bar{\partial}_{j]}L + y^m\partial_m g_{ij} - y^m N^k_m \bar{\partial}_i g_{kj}\right) && \text{(D.5)}\\ &= \tfrac{1}{2}\left(\partial_{[i}\bar{\partial}_{j]}L + y^m\partial_m g_{ij} - y^m N^k_m \bar{\partial}_k g_{ij}\right) && \text{(D.6)}\\ &= \tfrac{1}{2}\left(\partial_{[i}\bar{\partial}_{j]}L + y^m\delta_m g_{ij}\right). && \text{(D.7)} \end{aligned}$$

It thus follows that

$$N^k_i g_{kj} + N^k_j g_{ik} = y^k \delta_k g_{ij}, \tag{D.8}$$

which gives the desired result, proving the first implication.

Conversely, to prove that $(ii) \Rightarrow (i)$, note that we have

$$y^i \delta_i L = \boldsymbol{\nabla}_D L = \boldsymbol{\nabla}_D\left(g_{jk}y^j y^k\right) = \boldsymbol{\nabla}_D g_{jk} y^j y^k, \tag{D.9}$$

since $\boldsymbol{\nabla}_D y^i = 0$. So $\boldsymbol{\nabla}_D g_{ij}$ implies $0 = y^i\delta_i L$, which means also that

$$0 = \bar{\partial}_j\left(y^i\delta_i L\right) = \delta_j L + y^i \bar{\partial}_j \delta_i L. \tag{D.10}$$

Thus it suffices to show that $y^i\bar{\partial}_j\delta_i L = 0$. Since $[\bar{\partial}_j, \delta_i] = -\bar{\partial}_j N^k_i \bar{\partial}_k$ we have

$$\begin{aligned} y^i\bar{\partial}_j\delta_i L &= y^i\delta_i\bar{\partial}_j L - y^i\bar{\partial}_j N^k_i\bar{\partial}_k L && \text{(D.11)}\\ &= y^i\delta_i\bar{\partial}_j L - y^i\bar{\partial}_i N^k_j\bar{\partial}_k L - y^i T^k{}_{ij}\bar{\partial}_k L && \text{(D.12)}\\ &= y^i\delta_i\bar{\partial}_j L - N^k_j\bar{\partial}_k L - y^i T^k{}_{ij}\bar{\partial}_k L && \text{(D.13)}\\ &= \boldsymbol{\nabla}_D\left(\bar{\partial}_j L\right) - y^i T^k{}_{ij}\bar{\partial}_k L && \text{(D.14)}\\ &= \boldsymbol{\nabla}_D\left(2g_{jk}y^k\right) - y^i T^k{}_{ij}\bar{\partial}_k L && \text{(D.15)}\\ &= -y^i T^k{}_{ij}\bar{\partial}_k L && \text{(D.16)}\\ &= -2T_{kij}y^i y^k && \text{(D.17)} \end{aligned}$$

where in the second line we used the definition of the torsion tensor, in the third line we used homogeneity of N^k_{ij}, in the fourth line we used the definition of $\boldsymbol{\nabla}_D$, then we expressed $\bar{\partial}_j L$ in terms of g_{ij}, and finally in the last line we used the Leibniz rule for $\boldsymbol{\nabla}_D$ and the fact that $\boldsymbol{\nabla}_D y^k = 0$ and $\boldsymbol{\nabla}_D g_{jk} = 0$. Combining this

with what we found above, we conclude that if $\mathbf{\nabla}_D g = 0$ then

$$\delta_j L = y^i T^k{}_{ij} \bar{\partial}_k L \tag{D.18}$$

and hence if additionally the torsion of $\mathbf{\nabla}_D$ vanishes then we must indeed have $\delta_i L = 0$. □

D.2 Generality of generalized m-Kropina metrics

Theorem 5.3.12 assumes that $\phi' \neq 0$ on $\mathcal{A}$. Here we generalize the theorem to the case where no such condition is presupposed. In what follows, we will denote by s both the function $s : (x, y) \mapsto \beta(x, y)/\alpha(x, y)$ as well as the function values $s \in \mathbb{R}$. So we will write, for instance, that $s \in s(U)$ if the *value* s lies in the image of the set U under the *function* s. We start with a stronger version of Lemma 5.3.11, where the assumption that U is connected is omitted.

Lemma D.2.1. *Suppose that b_i is nowhere vanishing. If $U \subset \mathcal{A}$ is open then the image $s(U)$ is a union of intervals that each have nonempty interior, and $s(U) \subset \overline{s(U)^0 \setminus C}$ for any finite set C.*

Proof. Since U is an open subset of the manifold TM, it is itself a manifold and hence it is locally connected. As a result, we may write $U = \cup_i U_i$ where each U_i is open and connected. By Lemma 5.3.11, each $s(U_i)$ is an interval with nonempty interior, and we have $s(U_i) \subset \overline{s(U_i)^0 \setminus C}$ for each i. Hence

$$s(U) = \cup_i s(U_i) \subset \cup_i \overline{s(U_i)^0 \setminus C} \subset \overline{\cup_i s(U_i)^0 \setminus C} \subset \overline{(\cup_i s(U_i))^0 \setminus C} \tag{D.19}$$

$$= \overline{(s(\cup_i U_i))^0 \setminus C} = \overline{s(U)^0 \setminus C}, \tag{D.20}$$

as desired. □

As a result, if some ODE for $\phi(s)$ is satisfied for all $s \in s(U) \setminus C$, where $U \subset \mathcal{A}$ is some open set and C a finite set, then the ODE may be solved (explicitly) on the open set $s(U)^0 \setminus C$ using standard methods, and this uniquely determines ϕ on all of $s(U)$ by continuous extension. The following result is the generalization Theorem 5.3.12 that we alluded to.

Theorem D.2.2. *Let $F = \alpha\phi(\beta/\alpha)$ be a properly Finslerian (α, β)-metric with $\dim M > 2$. Suppose furthermore that $\mathcal{A}$ is connected. Then TFAE:*

(i) F is Berwald and there exist nowhere vanishing $\lambda, \rho, \sigma \in C^\infty(M)$ such that

$$T^k{}_{ij} = \lambda b^k b_i b_j + \rho\left(b_i \delta^k_j + b_j \delta^k_i\right) + \sigma b^k a_{ij}; \tag{D.21}$$

(ii) β is closed and there are nonvanishing constants c, d, m and a nowhere vanishing $p \in C^\infty(M)$ such that

$$\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2} \tag{D.22}$$

$$\nabla_i b_j = p\left\{cm|b|^2 a_{ij} + \left[c(1-m) + \epsilon d|b|^2\right] b_i b_j\right\} \tag{D.23}$$

In that case $\rho = -\sigma = -cmp$ and $\lambda = \epsilon dp$.

Proof. Assuming (ii) then, according to Proposition 5.3.9, F is Berwald and the affine connection is given by (5.40) and from this we infer that T^k_{ij} is clearly of the form (D.21) with $\lambda = \epsilon dp$ and $\rho = -\sigma = -cmp$. Hence (ii) implies (i). We will split the proof that (i) implies (ii) over the two lemmas below. Hence, with these lemmas, the proof of the theorem is complete. □

Denote $B = \{\phi' = 0\} \subset \mathcal{A}$ and $\tilde{\mathcal{A}} = \mathcal{A} \setminus \overline{B^0}$. There are two cases that we will distinguish. Either $B^0 = \emptyset$ or $B^0 \neq \emptyset$. Each of the two cases gets its own lemma.

Lemma D.2.3. *Theorem D.2.2 holds under the additional assumption that $B^0 = \emptyset$.*

Proof. It suffices to show that (i) implies (ii). The Berwald condition (5.11) must be satisfied on $\tilde{\mathcal{A}}$ with T^k_{ij} given by (D.21), in which case this condition reduces to

$$y^j \nabla_i b_j = \lambda|b|^2\beta b_i + 2\rho\beta b_i + \sigma|b|^2 y_i + \frac{\epsilon\psi}{\alpha}\left(\lambda\beta^2 b_i + \rho(\beta y_i + A b_i) + \sigma\beta y_i\right), \tag{D.24}$$

where $\psi = \phi/\phi' - s$. We can write this also as

$$y^j \nabla_i b_j - \left(\lambda|b|^2\beta b_i + 2\rho\beta b_i + \sigma|b|^2 y_i\right) = \frac{\epsilon\psi}{\alpha}\left(\lambda\beta^2 + \rho A\right) b_i + \frac{\epsilon\psi}{\alpha}(\rho+\sigma)\beta y_i, \tag{D.25}$$

where we have collected all manifestly linear terms on the LHS and grouped the RHS by a b_i term and a y_i term. Since the LHS is linear, the RHS must be so as well, and in fact, by Lemma B.0.2 this can only be achieved if both terms on the RHS are linear, i.e. if $\frac{\epsilon\psi}{\alpha}(\rho+\sigma)\beta$ is independent of y and if $\frac{\epsilon\psi}{\alpha}(\lambda\beta^2 + \rho A)$ is linear. Since $A = 0$ is excluded, by definition, from $\mathcal{A}$ and since $\mathcal{A}$ has only

a single connected component, by assumption, $\mathcal{A}$ must be contained in one of the open sets $A > 0$ or $A < 0$. And that, in particular, implies that $\epsilon = \operatorname{sgn}(A)$ is just a constant on $\mathcal{A}$. Thus we need $\frac{\psi}{\alpha}(\rho + \sigma)\beta$ to be independent of y and $\frac{\psi}{\alpha}(\lambda\beta^2 + \rho A)$ to be linear.

Step 1) Claim: $\rho = -\sigma$

First we consider the y_i term, which is linear in y iff $\psi(\rho + \sigma)\beta/\alpha =$ is independent of y. The only way this can be true, assuming F is properly Finslerian, is if $\rho = -\sigma$. To see this, suppose that $\rho \neq -\sigma$. Then we must have $\psi\beta/\alpha = \psi s =: \ell$ where ℓ is a priori a function of x, but since the LHS depends only on s, ℓ must actually be constant. Hence $\psi = \phi/\phi' - s = \ell/s$ for all $(x, y) \in \tilde{\mathcal{A}}$ with $s \neq 0$. As a differential equation in the variable s this must hold for all $s \in s(\tilde{\mathcal{A}}) \setminus \{0\}$. Since $\mathcal{A}$ is open and connected, we can apply Lemma D.2.1 and its consequences: we can integrate the differential equation for $\phi(s)$ on the interior of $s(\tilde{\mathcal{A}}) \setminus \{0\}$ and then extend the solution by continuity to all of $s(\tilde{\mathcal{A}})$, i.e. for all relevant values of s. Moreover, since B is assumed to have empty interior, $\tilde{\mathcal{A}}$ is dense in $\mathcal{A}$. Hence, by continuity, $s(\tilde{\mathcal{A}})$ is dense in $s(\mathcal{A})$. This means that ϕ can be obtained for *all* values of s by continuous extension.

We will now apply this scheme. Assuming wlog that $\ell \neq 0$ (for otherwise $\phi - s\phi' = 0$ and $\det g_{ij} = 0$, by (C.14)) we find $\ln|\phi| = \int (s + \ell/s)^{-1}\,\mathrm{d}s = \ln(|c(\ell + s^2)|)/2$, where c is an integration constant (that may, at this stage, be different on each connected component). Then $\phi = \sqrt{c(\ell + s^2)}$ for s in the interior of $s(\tilde{\mathcal{A}}) \setminus \{0\}$, and by the argument given above this must be the form of ϕ for all values of s. Hence $F^2 = \alpha^2\phi^2 = c(\ell\alpha^2 + \beta^2)$ is pseudo-Riemannian. This is excluded in the premise of the theorem and hence it is a contradiction, so we must indeed have $\rho = -\sigma$.

Step 2) Deriving an ODE in s

Next, we consider the b_i term, setting $\sigma = -\rho$. This term is linear iff the coefficient of b_i is linear, which is the case iff

$$\bar{\partial}_i\bar{\partial}_j\left[\frac{\psi}{\alpha}\left(\lambda\beta^2 + \rho A\right)\right] = 0. \tag{D.26}$$

For the first derivative, we have

$$\bar{\partial}_i \left[\frac{\psi}{\alpha} \left(\lambda\beta^2 + \rho A \right) \right] = \frac{\psi' \bar{\partial}_i s}{\alpha} \left(\lambda\beta^2 + \rho A \right) - \frac{\psi}{\alpha^2} \bar{\partial}_i \alpha \left(\lambda\beta^2 + \rho A \right) + \frac{\psi}{\alpha} \left(2\lambda\beta b_i + 2\rho y_i \right) \tag{D.27}$$

$$= \frac{\psi'}{\alpha^2} \left(b_i - \frac{\beta y_i}{A} \right) \left(\lambda\beta^2 + \rho A \right) - \frac{\epsilon \psi y_i}{\alpha^3} \left(\lambda\beta^2 + \rho A \right) + \frac{\psi}{\alpha} \left(2\lambda\beta b_i + 2\rho y_i \right) \tag{D.28}$$

$$= \frac{1}{\alpha^2} \left[- \left(\frac{\beta\psi'}{A} + \frac{\epsilon\psi}{\alpha} \right) \left(\lambda\beta^2 + \rho A \right) + 2\alpha\rho\psi \right] y_i + \{ b_i \text{ terms} \}, \tag{D.29}$$

where we have used the identities

$$\bar{\partial}_i \alpha = \frac{\epsilon y_i}{\alpha}, \qquad \bar{\partial}_i s = \frac{1}{\alpha} \left(b_i - \frac{\beta y_i}{A} \right), \tag{D.30}$$

and separated the y_i and the b_i parts, the latter of which are irrelevant. The reason for this is the following. By inspecting the situation (or working it out exactly) it is clear that the second derivative in (D.26) will be of the form $f a_{ij} + g h_{ij}$, where f, g are functions and h_{ij} is a linear combination of $b_i b_j, b_i y_j, y_i b_j$ and $y_i y_j$. Hence h_{ij} has at most rank 2, which is strictly less than the rank of a_{ij}. That means that (D.26), i.e. the equation $f a_{ij} + g h_{ij} = 0$, can only hold if $f = 0$, for otherwise we could write $a_{ij} = (g/f) h_{ij}$ and a_{ij} and h_{ij} would have the same rank. It turns out that it suffices for what we aim to prove to consider this coefficient f, as we will see shortly. Differentiation of the b_i terms in (D.29), however, will never yield terms proportional to a_{ij}. In fact, the only way we get terms proportional to a_{ij} is by directly differentiating y_i in (D.29). No other terms resulting from the product rule are proportional to a_{ij} either. This simplifies this calculation enormously because it means that in order to get the a_{ij}-coefficient f from (D.29) we effectively need only replace y_i by its derivative $\bar{\partial}_j y_i = a_{ij}$ to obtain

$$\bar{\partial}_i \bar{\partial}_j \left[\frac{\psi}{\alpha} \left(\lambda\beta^2 + \rho A \right) \right] = \tag{D.31}$$

$$\frac{1}{\alpha^2} \left[- \left(\frac{\beta\psi'}{A} + \frac{\epsilon\psi}{\alpha} \right) \left(\lambda\beta^2 + \rho A \right) + 2\alpha\rho\psi \right] a_{ij} + g h_{ij}. \tag{D.32}$$

As we argued above, the vanishing of this implies that we must have

$$0 = \frac{1}{\alpha}\left[-\left(\frac{\beta\psi'}{A}+\frac{\epsilon\psi}{\alpha}\right)\left(\lambda\beta^2+\rho A\right)+2\alpha\rho\psi\right] \tag{D.33}$$

$$= -\left(\frac{s\psi'}{\epsilon\alpha^2}+\frac{\epsilon\psi}{\alpha^2}\right)\left(\lambda\beta^2+\rho A\right)+2\rho\psi \tag{D.34}$$

$$= -\epsilon\left(s\psi'+\psi\right)\left(\lambda s^2+\rho\epsilon\right)+2\rho\psi, \tag{D.35}$$

or, after rewriting,

$$\frac{\psi'}{\psi} = \frac{\eta - s^2}{s(s^2+\eta)}, \qquad \eta \equiv \epsilon\rho/\lambda. \tag{D.36}$$

As before, since the LHS depends only on s, the RHS should also not depend on x explicitly, which implies that η, which is a priori a function on M, should actually be a constant. Hence (D.36) is an ODE for $\psi(s)$.

Step 3) Solving the ODE to obtain $\phi(s)$
In analogy with how we solved the differential equation for ϕ earlier in this proof, this one is solved uniquely by

$$\psi = \frac{\tilde{c}s}{s^2+\eta}, \tag{D.37}$$

for all $s \in s(\tilde{\mathcal{A}}) \setminus \{\sqrt{-\eta}, -\sqrt{-\eta}\}$, where $\tilde{c}$ is an integration constant. We find ϕ on K^0 via $\psi = \phi/\phi' - s$. This can now be written as

$$\frac{\phi'}{\phi} = \left(\frac{\tilde{c}s}{s^2+\eta}+s\right)^{-1} = \frac{ds^2-cm}{ds^3+cs}, \tag{D.38}$$

where we have introduced an arbitrary constant d and defined $c = d(\tilde{c}+\eta)$ and $m = -d\eta/c = -\eta/(\eta+\tilde{c})$. This is uniquely solved by

$$\phi(s) = \tilde{d}s^{-m}(c+ds^2)^{(m+1)/2}, \tag{D.39}$$

for $s \in s(\tilde{\mathcal{A}}) \setminus \{\sqrt{-\eta}, -\sqrt{-\eta}\}$, where we may absorb the integration constant $\tilde{d}$ (up to sign) into c and d. Hence ϕ attains the desired form

$$\phi(s) = \pm s^{-m}(c+ds^2)^{(m+1)/2} \tag{D.40}$$

for all $s \in s(\tilde{\mathcal{A}}) \setminus \{\sqrt{-\eta}, -\sqrt{-\eta}\}$. Once again, by Lemma D.2.1 and the fact that

$\tilde{\mathcal{A}}$ is dense in $\mathcal{A}$, we can again extend the solution by continuity to all relevant values of s. In other words, ϕ has the form (D.40) for all relevant values of s, as desired. In principle, the constants c, d, m and the sign $\pm$ could have different values on different connected components of $s(\mathcal{A})$, but since $\mathcal{A}$ is assumed to be connected, $s(\mathcal{A})$ is also connected and hence the constants and the sign are actually fixed over all of $s(\mathcal{A})$.

Step 4) The condition on $\nabla_i b_j$
Finally we substitute the form (D.37) of ψ and the relation between ρ, σ and λ into the Berwald condition (D.25), leading to

$$y^j \nabla_i b_j = \left(\lambda |b|^2 \beta b_i + 2\rho \beta b_i + \sigma |b|^2 y_i\right) + \epsilon \lambda \tilde{c} \beta b_i, \tag{D.41}$$

from which we can infer by differentiation that for $\nabla_i b_j$ is symmetric under $i \leftrightarrow j$. This implies that the 1-form β is closed and hence, by Proposition 5.3.9, that the desired Berwald condition holds. This completes the proof. □

Remark D.2.4. *As a consistency check for step 4 of the proof, we may also derive the desired Berwald condition directly from (D.41). For this, we note that the definitions of c and m in terms of $\tilde{c}$ and η above may be inverted as $\tilde{c} = c(1+m)/d$ and $\eta = -mc/d$. Using also the relations between λ, ρ, σ, we find that (D.25) turns into*

$$\begin{aligned}
y^j \nabla_i b_j &= \left(\lambda |b|^2 \beta b_i + 2\beta b_i + \sigma |b|^2 y_i\right) + \epsilon \lambda \tilde{c} \beta b_i && \text{(D.42)}\\
&= \left(\lambda |b|^2 \beta b_i + 2\epsilon\eta\lambda \beta b_i - \epsilon\eta\lambda |b|^2 y_i\right) + \epsilon\lambda \frac{c(1+m)}{d} \beta b_i && \text{(D.43)}\\
&= \lambda \left[\left(|b|^2 + 2\epsilon\eta + \epsilon \frac{c(1+m)}{d}\right) \beta b_i - \epsilon\eta |b|^2 y_i\right] && \text{(D.44)}\\
&= \lambda \left[\left(|b|^2 - \frac{2\epsilon mc}{d} + \frac{\epsilon c(1+m)}{d}\right) \beta b_i + \frac{\epsilon mc}{d} |b|^2 y_i\right] && \text{(D.45)}\\
&= \lambda \left[\left(|b|^2 + \frac{\epsilon c}{d}(1-m)\right) \beta b_i - \epsilon \frac{-mc}{d} |b|^2 y_i\right], && \text{(D.46)}
\end{aligned}$$

from which we can infer by differentiation that

$$\nabla_i b_j = \frac{\epsilon\lambda}{d} \left\{\left[\epsilon d |b|^2 + c(1-m)\right] b_i b_j + mc |b|^2 a_{ij}\right\}, \tag{D.47}$$

which is precisely (D.23) if we identify $p = \epsilon\lambda/d$.

Lemma D.2.5. *Theorem D.2.2 holds under the additional assumption that* $B^0 \neq \emptyset$.

Proof. Again, it suffices to prove that (i) implies (ii). Most of the proof of the previous lemma is still valid. In particular, the conclusion that ϕ must be of the form

$$\phi(s) = \pm s^{-m}(c + ds^2)^{(m+1)/2} \tag{D.48}$$

still holds on $s(\tilde{\mathcal{A}})$. The difference is that $\tilde{\mathcal{A}}$ is currently not dense in $\mathcal{A}$ and so we cannot extend this by continuity to all of $s(\mathcal{A})$. However, since we defined $B = \{\phi' = 0\} \subset \mathcal{A}$ and $\tilde{\mathcal{A}} = \mathcal{A} \setminus \overline{B^0}$, where we will agree that the closure is taken, by definition, in A, we have

$$\mathcal{A} = \tilde{\mathcal{A}} \cup \overline{B^0}, \qquad s(\mathcal{A}) = s(\tilde{\mathcal{A}}) \cup s(\overline{B^0}), \tag{D.49}$$

so it remains to see what happens for $s \in s(\overline{B^0})$. First, on $s(B^0)$, we have $\phi' = 0$, by definition. By Lemma D.2.1 and the remark right below it, this implies that

$$\phi(s) = k \tag{D.50}$$

for all $s \in s(\overline{B^0})$. Thus what we have shown is that for all s in $s(\mathcal{A})$, ϕ is given by either by (D.48) or by (D.50). Note that $s(\mathcal{A})$ is an interval because $\mathcal{A}$ is connected and ϕ assumed continuous, even smooth. Moreover, by Lemma D.2.1, both $s(\overline{B^0})$ and $s(\tilde{\mathcal{A}})$ are a union of intervals with nonempty interior. We want to prove that (D.48) holds *everywhere*. Suppose not. Consider the interface between an interval on which (D.48) is satisfied and an interval on which (D.50) is satisfied. ϕ is assumed to be smooth on this interface. That means that all derivatives of (D.48) must tend to zero (which are the derivatives of (D.50)) if s tends to this interface. However, one can check that there are no values of s for which all derivatives of (D.48) tend to zero, except if $cm = 0$, in which case F is either degenerate (which would be a contradiction) or pseudo-Riemannian. But that means that on every interval on which (D.48) is satisfied, F must in fact be pseudo-Riemannian. And that simply means that F is pseudo-Riemannian as a whole. This is a contradiction with our assumptions and therefore this completes the proof. □

List of Publications

This dissertation is based on:

[H1] C. Pfeifer, **S. Heefer**, A. Fuster. *Identifying Berwald Finsler Geometries*, Differential Geom. Appl. 79, 101817 (2021), arXiv:1909.05284

[H2] A. Fuster, **S. Heefer**, C. Pfeifer, N. Voicu. *On the Non Metrizability of Berwald Finsler Spacetimes*, Universe 6, 64 (2020), arXiv:2003.02300

[H3] **S. Heefer**, C. Pfeifer, J. van Voorthuizen, A. Fuster. *On the Metrizability of m-Kropina Spaces with Closed Null 1-form*, J. Math. Phys. 64, 022502 (2023), arXiv:2210.02718

[H4] **S. Heefer**, C. Pfeifer, A. Fuster. *Randers pp-Waves*, Phys. Rev. D 104, 024007 (2021), arXiv:2011.12969

[H5] **S. Heefer**, A. Fuster. *Finsler Gravitational Waves of (α, β)-type and their Observational Signature*, Class. Quantum Grav. 40 184002 (2023), arXiv:2302.08334

[H6] **S. Heefer**, C. Pfeifer, A. Reggio, A. Fuster. *A Cosmological Unicorn Solution to Finsler Gravity*, Phys. Rev. D 108, 064051 (2023), arXiv:2306.00722

Other journal articles:

A. Addazi *et al. Quantum Gravity Phenomenology at the Dawn of the Multi-Messenger Era – a Review*, Prog. Part. Nucl. Phys vol. 125, 103948 (2022), arXiv:2111.05659

G. Gubitosi, **S. Heefer**. *Relativistic Compatibility of the Interacting κ-Poincaré Model and Implications for the Relative Locality Framework*, Phys. Rev. D 99, 086019 (2019), arXiv:1903.04593

Preprints:

R. Alves Batista *et al. White Paper and Roadmap for Quantum Gravity Phenomenology in the Multi-Messenger Era*, arXiv:2312.00409 (submitted to *Class. Quantum Grav.*)

In preparation:

S. Heefer, L.F. Niehof, A. Fuster. *pp-Waves with Wave-Surfaces of Constant Curvature in Quadratic Gravity*

S. Heefer. *Local Metrizability and Ricci-Flatness of Berwald m-Kropina Spaces*

Bibliography

[1] A. Addazi, J. Alvarez-Muniz, *et al.*, "Quantum gravity phenomenology at the dawn of the multi-messenger era—a review," *Prog. Part. Nucl. Phys.* **125** (2022) 103948, arXiv:2111.05659.

[2] R. A. Batista, G. Amelino-Camelia, *et al.*, "White paper and roadmap for quantum gravity phenomenology in the multi-messenger era," arXiv:2312.00409.

[3] D. Mattingly, "Modern tests of Lorentz invariance," *Living Rev. Rel.* **8** (2005) 5, arXiv:gr-qc/0502097.

[4] S. Chern, "Finsler geometry is just Riemannian geometry without the quadratic restriction," *Notices Amer. Math. Soc.* **43** (1996) 959–963.

[5] B. Riemann, "Über die Hypothesen, welche der Geometrie zu Grunde liegen," *Abh. Königl. Ges. Wiss. Göttingen* **13** (1868) 133–150.

[6] B. Riemann, "On the hypotheses which lie at the bases of geometry," *Nature* **8** (1873) 14–17.

[7] P. Finsler, *Über Kurven und Flächen in allgemeinen Räumen*. PhD dissertation, Georg-August Universität zu Göttingen, 1918.

[8] F. Girelli, S. Liberati, and L. Sindoni, "Planck-scale modified dispersion relations and Finsler geometry," *Phys. Rev. D* **75** (2007) 064015, arXiv:gr-qc/0611024.

[9] D. Raetzel, S. Rivera, and F. P. Schuller, "Geometry of physical dispersion relations," *Phys. Rev. D* **83** (2011) 044047, arXiv:1010.1369.

[10] E. Rodrigues and I. P. Lobo, "Revisiting Legendre transformations in Finsler geometry," arXiv:2208.11406. (2022).

[11] G. Amelino-Camelia, L. Barcaroli, G. Gubitosi, S. Liberati, and N. Loret, "Realization of doubly special relativistic symmetries in Finsler geometries," *Phys. Rev. D* **90** (2014) 125030, arXiv:1407.8143.

[12] I. P. Lobo, N. Loret, and F. Nettel, "Investigation of Finsler geometry as a generalization to curved spacetime of Planck-scale-deformed relativity in the de Sitter case," *Phys. Rev. D* **95** (2017) 046015, arXiv:1611.04995.

[13] M. Letizia and S. Liberati, "Deformed relativity symmetries and the local structure of spacetime," *Phys. Rev. D* **95** (2017) 046007, arXiv:1612.03065.

[14] R. K. Tavakol and N. Van Den Bergh, "Finsler spaces and the underlying geometry of space-time," *Phys. Lett. A* **112** (1985) 23–25.

[15] R. K. Tavakol and N. Van den Bergh, "Viability criteria for the theories of gravity and Finsler spaces," *Gen. Relativ. Gravit.* **18** (1986) 849–859.

[16] R. Tavakol, "Geometry of spacetime and Finsler geometry," *Int. J. Mod. Phys. A* **24** (2009) 1678–1685.

[17] J. Ehlers, F. A. E. Pirani, and A. Schild, "Republication of: The geometry of free fall and light propagation," *Gen. Relativ. Gravit.* **44** (2012) 1587–1609.

[18] C. Lämmerzahl and V. Perlick, "Finsler geometry as a model for relativistic gravity," *Int. J. Geom. Methods Mod.* **15** (2018) 1850166, arXiv:1802.10043.

[19] A. N. Bernal, M. A. Javaloyes, and M. Sánchez, "Foundations of Finsler spacetimes from the observers' viewpoint," *Universe* **6** (2020) 55, arXiv:2003.00455.

[20] L. Bubuianu and S. I. Vacaru, "Axiomatic formulations of modified gravity theories with nonlinear dispersion relations and Finsler–Lagrange–Hamilton geometry," *The Eur. Phys. J. C* **78** (2018) 969.

[21] V. Perlick, "Characterization of standard clocks by means of light rays and freely falling particles," *Gen. Relativ. Gravit.* **19** (1987) 1059–1073.

[22] H. F. M. Goenner, "On the history of unified field theories," *Living Rev. Rel.* **7** (2004) 2.

[23] H. F. M. Goenner, "On the history of unified field theories. Part II. (ca. 1930–ca. 1965)," *Living Rev. Rel.* **17** (2014) 5.

[24] M. Tonnelat, *Einstein's Theory of Unified Fields*. Routledge Library Editions: 20th Century Science. Taylor & Francis, 2014.

[25] J. K. Beem, "Indefinite Finsler spaces and timelike spaces," *Can. J. Math.* **22** (1970) 1035.

[26] G. S. Asanov, *Finsler Geometry, Relativity and Gauge Theories*. D. Reidel Publishing Company, 1985.

[27] S. Rutz, "A Finsler generalisation of Einstein's vacuum field equations," *Gen. Relativ. Gravit.* **25** (1993) 1139.

[28] C. Pfeifer and M. N. R. Wohlfarth, "Finsler geometric extension of Einstein gravity," *Phys. Rev. D* **85** (2012) 064009, arXiv:1112.5641.

[29] J. I. Horváth, "A geometrical model for the unified theory of physical fields," *Phys. Rev.* **80** (1950) 901–901.

[30] J. I. Horváth and A. Moór, "Entwicklung einer einheitlichen Feldtheorie begründet auf die Finslersche Geometrie," *Zeitschrift für Physik* **131** (1952) 544–570.

[31] S. Ikeda, "On the theory of fields in Finsler spaces," *J. Math. Phys.* **22** (1981) 1215–1218.

[32] Z. Chang and X. Li, "Modified Friedmann model in Randers-Finsler space of approximate Berwald type as a possible alternative to dark energy hypothesis," *Phys. Lett.* **B676** (2009) 173–176, arXiv:0901.1023.

[33] A. P. Kouretsis, M. Stathakopoulos, and P. C. Stavrinos, "The general very special relativity in Finsler cosmology," *Phys. Rev. D* **79** (2009) 104011, arXiv:0810.3267.

[34] P. Stavrinos, O. Vacaru, and S. I. Vacaru, "Modified Einstein and Finsler like theories on tangent Lorentz bundles," *Int. J. Mod. Phys. D* **23** (2014) 1450094, arXiv:1401.2879.

[35] N. Voicu, "New considerations on Hilbert action and Einstein equations in anisotropic spaces," *AIP Conf. Proc.* **1283** (2010) 249–257, arXiv:0911.5034.

[36] E. Minguzzi, "The connections of pseudo-Finsler spaces," *Int. J. Geom. Meth. Mod. Phys.* **11** (2014) 1460025, arXiv:1405.0645. [Erratum: Int. J. Geom. Meth. Mod. Phys. 12, no.7, 1592001 (2015)].

[37] C. Pfeifer, *The Finsler spacetime framework: backgrounds for physics beyond metric geometry*. PhD dissertation, University of Hamburg, 2013.

[38] M. Hohmann, C. Pfeifer, and N. Voicu, "Finsler gravity action from variational completion," *Phys. Rev. D* **100** (2019) 064035, arXiv:1812.11161.

[39] N. Voicu and D. Krupka, "Canonical variational completion of differential equations," *J. Math. Phys.* **56** (2015) 043507, arXiv:1406.6646.

[40] M. Á. Javaloyes, M. Sánchez, and F. F. Villaseñor, "The Einstein-Hilbert-Palatini formalism in pseudo-Finsler geometry," arXiv:2108.03197. (2023). To appear in Adv. Theor. Math. Phys.

[41] J. Szilasi, R. L. Lovas, and D. Cs. Kertész, "Several ways to Berwald manifolds - and some steps beyond," *Extracta Math.* **26** (2011) 89–130, arXiv:1106.2223.

[42] Z. Szabó, "Positive definite Berwald spaces," *Tensor* **35** (1981) 25–39.

[43] A. Fuster and C. Pabst, "Finsler *pp*-waves," *Phys. Rev. D* **94** (2016) 104072, arXiv:1510.03058.

[44] A. Fuster, C. Pabst, and C. Pfeifer, "Berwald spacetimes and very special relativity," *Phys. Rev. D* **98** (2018) 084062, arXiv:1804.09727.

[45] M. Rakhmanov, "On the round-trip time for a photon propagating in the field of a plane gravitational wave," *Class. Quantum Gravity* **26** (2009) 155010, arXiv:1407.5376.

[46] D. Bao, "On two curvature-driven problems in Riemann-Finsler geometry," in *Finsler Geometry, Sapporo 2005: In Memory of Makoto Matsumoto*, vol. 48 of *Adv. Stud. Pure Math.*, pp. 19–71. Mathematical Society of Japan, 2007.

[47] G. Asanov, "Finsleroid-Finsler spaces of positive-definite and relativistic types," *Rep. Math. Phys.* **58** (2006) 275–300.

[48] Z. Shen, "On a class of Landsberg metrics in Finsler geometry," *Can. J. Math.* **61** (2009) 1357–1374.

[49] S. Elgendi, "Solutions for the Landsberg unicorn problem in Finsler geometry," *J. Geom. Phys.* **159** (2021) 103918, arXiv:1908.10910.

[50] J. M. Lee, *Introduction to Riemannian Manifolds*. Graduate texts in mathematics. Springer Cham, 2018.

[51] B. O'Neill, *Semi-Riemannian Geometry With Applications to Relativity*. ISSN. Elsevier Science, 1983.

[52] A. Sasane, *A Mathematical Introduction To General Relativity*. World Scientific Publishing Company, 2021.

[53] L. W. Tu, *Differential Geometry: Connections, Curvature, and Characteristic Classes*. Graduate Texts in Mathematics. Springer Cham, 2017.

[54] L. W. Tu, *An Introduction to Manifolds*. Springer New York, 2010.

[55] J. Lee, *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer New York, 2012.

[56] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann-Finsler geometry*. Springer New York, 2000.

[57] J. Szilasi, *Connections, Sprays and Finsler Structures*. World Scientific, 2014.

[58] R. Miron and I. Bucataru, *Finsler Lagrange geometry: Applications to Dynamical Systems*. Editura Academiei Romane, 2007.

[59] Z. Shen, *Lectures on Finsler Geometry*. Series on Multivariate Analysis Series. World Scientific, 2001.

[60] P. L. Antonelli, R. S. Ingarden, and M. Matsumoto, *The Theory of Sprays and Finsler Spaces with Applications in Physics and Biology*. Kluwer Academic Publishers, 1993.

[61] P. Antonelli *et al.*, *Handbook of Finsler geometry, Vol. 1*. Handbook of Finsler Geometry. Kluwer Academic Publishers, 2003.

[62] P. Antonelli *et al.*, *Handbook of Finsler Geometry, Vol. 2*. Handbook of Finsler Geometry. Kluwer Academic Publishers, 2003.

[63] Z. Shen, *Differential Geometry of Spray and Finsler Spaces*. Springer Dordrecht, 2001.

[64] Y. Shen and Z. Shen, *Introduction To Modern Finsler Geometry*. World Scientific Publishing Company, 2016.

[65] D. Bao and C. Robles, "Ricci and flag curvatures in Finsler geometry," in *A Sampler of Riemann-Finsler Geometry*, vol. 50 of *Math. Sci. Res. Inst. Publ.*, pp. 197–259. Cambridge University Press, 2004.

[66] M. Anastasiei, "Finsler vector bundles – metrizable connections," *Period. Math. Hung.* **48** (2004) 83–91.

[67] M. Anastasiei, "Finsler connections in generalized Lagrange spaces," *Balk. J. Geom. its Appl.* **1** (1996) 1–9.

[68] J. Grifone, "Structure presque tangente et connexions I," *Ann. de l'Institut Fourier* **22** (1972) 287–334.

[69] M. Á. Javaloyes, M. Sánchez, and F. F. Villaseñor, "Anisotropic connections and parallel transport in Finsler spacetimes," in *Developments in Lorentzian Geometry*, A. L. Albujer, M. Caballero, A. García-Parrado, J. Herrera, and R. Rubio, eds., pp. 175–206. Springer International Publishing, Cham, 2022. arXiv:2107.05986.

[70] A. Tayebi, "A survey on unicorns in Finsler geometry," *AUT J. Math. Comput.* **2** (2021) 239–250.

[71] M. Hohmann, C. Pfeifer, and N. Voicu, "Cosmological Finsler spacetimes," *Universe* **6** (2020) 65, arXiv:2003.02299.

[72] N. Voicu, *Geometric Methods of Finsler-Based Field Theory*. Habilitation thesis, Transilvania University of Brasov, 2022.

[73] M. Matsumoto, *Finsler Geometry in the 20th-Century (Published in Handbook of Finsler Geometry, edited by P. L. Antonelli)*. Kluwer Academic Publishers, Dordrecht, 2003.

[74] S. Lang, *Algebra*. Springer New York, 2012.

[75] G. Randers, "On an asymmetrical metric in the four-space of general relativity," *Phys. Rev.* **59** (1941) 195–199.

[76] V. Kropina, "On projective two-dimensional Finsler spaces with special metric," *Trudy Sem. Vektor. Tenzor. Anal.* **11** (2006) 277–292, arXiv:math/0605759.

[77] W. Kundt, "The plane-fronted gravitational waves," *Zeitschrift für Physik* **163** (1961) 77–86.

[78] W. Kundt, G. J. Whitrow, H. Bondi, D. Bohm, W. B. Bonnor, A. Trautman, J. L. Synge, and A. Lichnerowicz, "Exact solutions of the field equations: twist-free pure radiation fields," *Proc. Roy. Soc. Lond. A* **270** (1962) 328–334.

[79] H. Stephani, D. Kramer, M. MacCallum, C. Hoenselaers, and E. Herlt, *Exact Solutions of Einstein's Field Equations*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2 ed., 2003.

[80] C. Vincze, "A new proof of Szabó's theorem on the Riemann-metrizability of Berwald manifolds," *Acta Math. Paedagog. Nyházi.* **21** (2005) 199–204.

[81] M. Crampin, "On the construction of Riemannian metrics for Berwald spaces by averaging," *Houston journal of mathematics* **40** (2014) 737–750.

[82] S. Chern, W. Chen, and K. Lam, *Lectures on Differential Geometry*. Series on university mathematics. World Scientific, 1999.

[83] V. Pravda, A. Pravdová, A. Coley, and R. Milson, "All spacetimes with vanishing curvature invariants," *Class. Quantum Gravity* **19** (2002) 6213, arXiv:gr-qc/0209024.

[84] Podolský, J. and Steinbauer, R. and Švarc, R., "Gyratonic pp-waves and their impulsive limit," *Phys. Rev. D* **90** (2014) 044050, arXiv:1406.3227.

[85] J. W. Maluf, J. F. da Rocha-Neto, S. C. Ulhoa, and F. L. Carneiro, "Kinetic energy and angular momentum of free particles in the gyratonic pp-waves space-times," *Class. Quantum Gravity* **35** (2018) 115001, arXiv:1801.04957.

[86] J. Ehlers and W. Kundt, "Exact solutions of the gravitational field equations," in *Gravitation: An Introduction to current research*, L. Witten, ed., pp. 49–101. John Wiley & Sons, Inc., 1962.

[87] C. Pfeifer, "Finsler spacetime geometry in physics," *Int. J. Geom. Meth. Mod. Phys.* **16** (2019) 1941004, arXiv:1903.10185.

[88] M. Hohmann, C. Pfeifer, and N. Voicu, "Relativistic kinetic gases as direct sources of gravity," *Phys. Rev. D* **101** (2020) 024062, arXiv:1910.14044.

[89] M. Hohmann, C. Pfeifer, and N. Voicu, "The kinetic gas universe," *Eur. Phys. J. C* **80** (2020) 809, arXiv:2005.13561.

[90] H. Andréasson, "The Einstein-Vlasov System/Kinetic Theory," *Living Rev. Rel.* **14** (2011) 4, arXiv:1106.1367.

[91] O. Sarbach and T. Zannias, "Relativistic kinetic theory: An introduction," *AIP Conf. Proc.* **1548** (2013) 134–155, arXiv:1303.2899.

[92] R. Hama, T. Harko, and S. V. Sabau, "Dark energy and accelerating cosmological evolution from osculating Barthel–Kropina geometry," *The Eur. Phys. J. C* **82** (2022) 385, arXiv:2204.04506.

[93] Y. Tong, "General teleparallel gravity from Finsler geometry," *Commun. Theor. Phys.* **75** (2023) 095403.

[94] C. Pfeifer and M. N. R. Wohlfarth, "Causal structure and electrodynamics on Finsler spacetimes," *Phys. Rev. D* **84** (2011) 044039, arXiv:1104.1079.

[95] C. Lämmerzahl, V. Perlick, and W. Hasse, "Observable effects in a class of spherically symmetric static Finsler spacetimes," *Phys. Rev. D* **D86** (2012) 104042, arXiv:1208.0619.

[96] M. A. Javaloyes and M. Sánchez, "Finsler metrics and relativistic spacetimes," *Int. J. Geom. Methods Mod. Phys.* **11** (2014) 1460032, arXiv:1311.4770.

[97] M. A. Javaloyes and M. Sánchez, "On the definition and examples of cones and Finsler spacetimes," *RACSAM* **114** (2019) 30, arXiv:1805.06978.

[98] G. S. Asanov, "Gravitational field equations based on Finsler geometry," *Found. Phys.* **13** (1983) 501–527.

[99] K. Landsman, *Foundations of General Relativity*. Radboud University Press, 2021.

[100] B. Chen and Y.-B. Shen, "On a class of critical Riemann-Finsler metrics," *Publ. Math. Debrecen* **72/3-4** (2008) 451–468.

[101] J. A. Viaclovsky, "Critical metrics for Riemannian curvature functionals," arXiv:1405.6080. (2014).

[102] N. Voicu, C. Pfeifer, and S. Cheraghchi, "Birkhoff theorem for Berwald Finsler spacetimes," arXiv:2306.07866. (2023). To appear in Phys. Rev. D.

[103] P. Jordan, J. Ehlers, and W. Kundt, "Strenge Lösungen der Feldgleichungen der Allgemeinen Relativitätstheorie," *Akad. Wiss. Lit. Mainz. Abhandl., Math.-Nat. Kl.* **2** (1960) 21–105.

[104] G. S. Hall and D. P. Lonie, "Holonomy groups and spacetimes," *Class. Quantum Gravity* **17** (2000) 1369–1382, arXiv:gr-qc/0310076.

[105] C. Batista, "Killing-Yano tensors of order $n-1$," *Class. Quantum Gravity* **31** (2014) 165019, arXiv:1405.4148.

[106] H. W. Brinkmann, "Einstein spaces which are mapped conformally on each other," *Math. Ann.* **94** (1925) 119–145.

[107] N. Rosen, "Plane polarized waves in the general theory of relativity," *Phys. Z. Sowjetunion* **12** (1937) 366–372.

[108] M. Blau, "Plane waves and Penrose limits," http://www.blau.itp.unibe.ch/lecturesPP.pdf, 2011.

[109] M. Blau and M. O'Loughlin, "Homogeneous plane waves," *Nucl. Phys. B.* **654** (2003) 135–176, arXiv:hep-th/0212135.

[110] S. S. Chern and Z. Shen, *Riemann-Finsler Geometry*. World Scientific, 2005.

[111] N. Voicu, A. Friedl-Szász, E. Popovici-Popescu, and C. Pfeifer, "The Finsler spacetime condition for (α, β)-metrics and their isometries," *Universe* **9** (2023) 198, arXiv:2302.09937.

[112] R. Bhatia, *Matrix Analysis*. Springer New York, 1997.

[113] A. G. Cohen and S. L. Glashow, "Very special relativity," *Phys. Rev. Lett.* **97** (2006) 021601, arXiv:hep-ph/0601236.

[114] G. Gibbons, J. Gomis, and C. Pope, "General very special relativity is Finsler geometry," *Phys. Rev. D* **76** (2007) 081701, arXiv:0707.2174.

[115] G. Yu. Bogoslovsky, "On a special relativistic theory of anisotropic space-time," *Proc. USSR Acad. Sci.* **213** (1973) 1055–1058.

[116] G. Yu. Bogoslovsky, "A special-relativistic theory of the locally anisotropic space-time," *Nuov. Cim. B* **40** (1977) 99.

[117] D. Harville, *Matrix Algebra From a Statistician's Perspective*. Springer New York, 1997.

[118] V. Perlick, "On the radar method in general-relativistic spacetimes," in *Lasers, Clocks and Drag-Free Control: Exploration of Relativistic Gravity in Space*, H. Dittus, C. Lämmerzahl, and S. G. Turyshev, eds., pp. 131–152. Springer Berlin Heidelberg, 2008.

[119] C. Pfeifer, "Radar orthogonality and radar length in Finsler and metric spacetime geometry," *Phys. Rev. D* **90** (2014) 064052, arXiv:1408.5306.

[120] N. Gürlebeck and C. Pfeifer, "Observers' measurements in premetric electrodynamics: time and radar length," *Phys. Rev. D* **97** (2018) 084043, arXiv:1801.07724.

[121] S. Cheraghchi, C. Pfeifer, and N. Voicu, "Four-dimensional SO(3)-spherically symmetric Berwald Finsler spaces," *Int. J. Geom. Methods Mod.* **20** (2022) 2350190, arXiv:2212.08603.

[122] M. Hohmann, C. Pfeifer, and N. Voicu, "Mathematical foundations for field theories on Finsler spacetimes," *J. Math. Phys.* **63** (2022) 032503, arXiv:2106.14965.

[123] S. Elgendi, "On the problem of non-Berwaldian Landsberg spaces," *Bull. Aust. Math. Soc.* **102** (2020) 331–341, arXiv:1801.08763.

[124] N. Straumann, *General Relativity*. Graduate Texts in Physics. Springer Dordrecht, 2012.

[125] P. Jordan, J. Ehlers, and W. Kundt, "Republication of: Exact solutions of the field equations of the general theory of relativity," *Gen. Relativ. Gravit.* **41** (2009) 2191–2280.

Acknowledgments

The past couple of years have been an incredible time for me. It has been a blessing to have had the opportunity to completely immerse myself in the world of academia, work on such fascinating topics, and go to numerous inspiring conferences all around Europe. There are quite some people who have been instrumental in this journey and I would like to express my sincere gratitude to all of them.

First of all, my co-promotor and daily supervisor, Andrea, thanks to whom all of this has been possible. Your research ideas and advice provided a solid foundation from the outset. You also gave me the freedom to pursue my own ideas without too much pressure and never forced me to do anything I was not comfortable with (except maybe take some days off from time to time—to my own benefit, of course). This has resulted in a very comfortable and productive way of working and doing research.

Second, my promotor, Luc, who would always ask good and difficult questions about my research that would incentivize me to think more deeply about certain aspects of it. Thanks for the support and the great discussions.

All other people who I collaborated and coauthored with over the last couple of years, most notably Christian and Nicoleta. I thoroughly enjoyed our joint projects and I hope we will continue our collaboration in the future.

All people—too many to name—with whom I had inspiring conversations and discussions. I'm particularly grateful for having been part of the COST action CA18108: *Quantum gravity phenomenology in the multi-messenger approach (QG-MM)*, which considerably broadened my perspective not just on my own research but on science in general.

My committee members. Thanks, first of all, for taking the time to read my dissertation and join the defense ceremony. But I'm especially thankful for the valuable feedback that I received on the draft version, which has certainly improved the quality of the dissertation.

My office mates and fellow ADG group members, Andrea, Bart, Finn, Gautam, Gijs, Georgios, Lars, Luc, Nicky, Remco, Rick, Stephan. I greatly enjoy our discussions (about math or otherwise) over lunch, coffee or just when we really ought to be working. I'm sure we'll have many more. Special thanks to Nicky and Rick for being my paranymphs. And special thanks to Gijs for taking the time to proofread the dissertation—in quite some technical detail. Thanks also to all my other colleagues at TU/e (CASA and beyond) who have made for a very pleasant atmosphere at the university.

I warmly thank all my family and friends for their interest in my endeavors, their support, and the welcome distractions. In particular, a big thanks goes out to the notorious 'Vrijdagavond' group, Dirk, Frank, Ivo, Jasper, Joran, Kjell, Loek, Niek, Roeland, Thomas, Tim, Tom, and Thijs, with whom I have so many fantastic memories. I count myself very fortunate to have such a great group of friends. Special thanks to Loek for taking the time to proofread a part of the dissertation.

Finally and perhaps most importantly, I'm deeply thankful to my parents and my brother, who have always supported me unconditionally and believed in me. Thank you for always being there for me.

Sjors Heefer
Eindhoven, April 2024

About the Author

Sjors Heefer was born on February 24, 1994 in 's-Hertogenbosch, the Netherlands. After finishing his Dutch VWO high school degree in 2012 at the ds. Pierson College in 's-Hertogenbosch, he went on to study Physics and Astronomy at Radboud University, Nijmegen, the Netherlands. Here he obtained his BSc degree in 2016 and his MSc degree in 2019, specializing in mathematical physics and graduating within the quantum gravity group with a master's thesis titled: *Advancements on the Covariance of the κ-Poincaré Model and Relative Locality.* In 2018 he started his PhD-TA at Eindhoven University of Technology, Department of Mathematics and Computer Science; a special five-year position combining PhD research with a considerable amount of teaching responsibility. His research has mainly focused on Finsler geometry and its application to gravity and spacetime. The results are presented in this dissertation. During his time as a PhD student, Sjors also obtained his university teaching qualification (UTQ). Moreover, after having been nominated twice, he won the award *Best teacher Applied Physics year 2* for his lectures given for the course *General Relavitity.* As of November 2023, he has been employed as a lecturer at the Department of Mathematics and Computer Science, Eindhoven University of Technology.